

An exploratory analysis of public health communication interventions and their effect on GP referral rates

Andrew White
BSc (Hons)



A thesis submitted in partial fulfilment of the requirements for the degree of
Master of Research
Bournemouth University
Faculty of Media and Communication / Science and Technology
October 2021

Abstract

Decreases in the number of GPs, increase in per-patient consultations and ageing populations are placing growing pressure on primary healthcare services around the world. At scale, effective public health initiatives are seen as a 'prevention is better than cure' solution to the long-term mitigation of these challenges.

This study reviews how referral rates from general practitioners are impacted through communication interventions in interprofessional healthcare settings. The investigation is conducted using a case study analysis of a public health authority organisation Public Health Dorset. Drawing upon their implementation and existing data regarding referrals to public health entities, this research reviews the organisation's effectiveness in this area. Through means of case-study research, a chronology of archival public health communication interventions was captured. The effectiveness of the interventions was quantified by utilising a causal inference analysis of the trend in the number of referrals following targeted communication activities over a 37 month period.

The results did not demonstrate a statistical significance in a change in the rate of GP referrals to the target health programme. This outcome was, in part, attributed to systematically poor tracking of intervention delivery. Supporting analysis identified seasonality trends as affecting both the rate of referrals from GPs and alternative sources such as self-referral. These trends suggested greater importance of patient participation in decision making as part of the referral process. Recommendations are proposed for the delivery, supervision and analysis of communication interventions at an organisational level in a primary healthcare setting.

Data analysis written in Python and submitted supporting this thesis in part of the fulfilment requirements for the degree of Master of Research can be found at the following website location <https://andyist.github.io/mres/>

Acknowledgements

The texts contained within this document are only possible thanks to the wide range of support I received during the project. I would, therefore, like to express my sincere gratitude to my advisors; Prof. Einar Thorsen for the support, experience and encouragement through the project, and similarly, Dr. Anna Feigenbaum for continued motivation, and insightful feedback which contributed to my own development alongside the work.

This research was originally envisaged and facilitated by Dr. Chris Skelly and Chris Rickets, my thanks goes to them and their colleagues at Public Health Dorset who provided insight and expertise that assisted this study as it progressed, and not least to Matt Freeman for his help during material gathering, and Darryl Houghton for assisting with the all-important archival data acquisition.

Finally, friends and family who's words of encouragement along the way, and I offer a very special thanks to my work colleagues and their support of my desire to step away from the company to conduct this research. Without them, this entire endeavour would simply have not been possible.

Funding

This research project was funded by Public Health Dorset.

Contents

Abstract	1
Acknowledgements	2
Funding	2
Contents	3
1. Introduction	5
1.1 Motivation	7
1.2 The Problem	8
1.3 Scope	9
1.4 Outline	11
2. Literature Review	13
2.1 Introduction	13
2.2 Professional Medical Communication	13
2.2.1 Background	13
2.2.2 Attributes and outcomes	15
2.2.3 Inter-organisational communication	17
2.2.4 Summary	21
3. Research Methodology	23
3.1 Introduction	23
3.2 Research Strategy	25
3.3 Research Approach	27
3.3.1 Identifying attributes of professional healthcare communication	27
3.3.2 Chronology of Case Study Communication Activities	28
3.3.3 Quantifying Service Referrals	29
3.3.4 Inferring causation	34
3.5 Ethical Considerations & Risks	35
3.6 Assumption	36
4. Results	37
4.1 Introduction	37
4.2 Intervention Chronology	37
4.3 Referrals	38

4.3.1 GP and Non-GP Referral rates	38
4.3.2 Client Demographics	42
4.3.3 Case Outcomes	43
4.3.4 Practice referrals	45
5. Discussion	47
5.1 Chronology and causal inference	47
5.2 Referrals	50
5.2.1 Absolute referrals by month	51
5.2.2 Referral Trend analysis	52
5.2.3 Demographics	53
5.2.4 Case Outcomes	53
6. Conclusion	54
6.1 RQ1 & RQ2	54
6.2 Limitations	56
6.3 Future Directions	57
6.4 Conclusion Summary	58
References	60
Appendixes	66
Appendix A	66
Appendix B	68
Appendix C	69
Appendix D	70
Appendix E	70

1. Introduction

The general health and well-being of a population places direct demand upon national healthcare services, which operate predominantly in a reactive capacity. Healthcare models vary around the world, this study explores the UK setting, in particular that of primary care's role in support of preventative public health initiatives.

Primary care is understood to include first-contact medical care provided continuously by the same doctor to individual patients (Onion and Berrington, 1999). Hogg et al., (2008) suggest that other important aspects of primary care include patient-provider relationships as defined by communication, holistic care and an awareness of the patient's family and culture. "The role of general practice is to act as the gatekeeper to the NHS, managing the health care for the practice population and referring relevant cases to secondary care services" (Peckham et al., 2015). In the UK, these services are responsible for delivering healthcare to a population of circa 54.3 million people (NHS, 2016). With over 300 million consultations occurring in general practice in England each year, General practitioners (GPs), or primary care physicians, have opportunities to provide advice, brief interventions and referral to targeted services. Therefore, they are recognized as playing an important role in promoting healthy lifestyles at individual patient encounters (Onion and Berrington, 1999). The Royal College of General Practitioners argues that GPs should be proactive in conducting public health activities and interventions, and it is expected that GPs should possess a wide range of skills related to ill-health prevention and public health (Peckham et al., 2015)

Placing GPs as the focus, is to look at their relatively unique position as an authoritative interface to the public they serve, not just to provide medical care but also to promote the health and well-being of the practice population" (Peckham et al., 2015). In NHS England's GP patient survey (2017) "More than nine out of ten patients (91.9%) have confidence and trust in the last GP they saw. This trust and confidence is a favourable resource when attempting to offer advice to the public about more general health and wellbeing. Taking this point further, it can be considered that the public with which they deal regularly are those often in need of improved wellbeing - acute conditions regularly compounded by unhealthy lifestyles. There has been a clear shift in thinking about practitioner roles that increasingly highlights a health-promoting role and one that also emphasises the practitioner's role as a facilitator as well as medical or health professional expert (Peckham et al, 2015). Since January 2012, there has been a call to NHS professionals to "Make every contact count", with staff using every contact they have with people as an "opportunity to help people stay in good health – by not smoking, eating healthily, drinking less alcohol, and exercising more." (NHS Future Forum. Summary Report, 2012 (p.8)), (Department of Health, 2012).

Policy support for public health appeared in the GP contract in 1990, which introduced the first payments for specific preventative care activities. The 2010–15 coalition government's programme published referenced proposals, including specific incentives for GPs to tackle public health problems (Wanless, 2004). In light of these circumstances, this study considered how GPs themselves might be motivated in the task of facilitating their patient uptake of public health programmes.

Organisations such as public health authorities, clinical commissioning groups, NHS trusts, various supporting healthcare clinics and services, all share the common goal of improved patient outcomes, yet are functionally different. Therefore, communication between them, which succeeds in positive patient care carries the greatest relevance. With a large variety of organisations complexity in service delivery is to be expected, Boon (2007) categorises the interchange and interrelationships in these circumstances as 'complex healthcare systems' pointing to the delivery of outcomes as having "an inherent self-organizing property and that the elements of complex systems themselves interact in such a way that through the interplay of the elements new properties emerge that cannot be seen when investigating only the component parts."

The identification of evidence of effective communication was seen to inform ongoing improvement and best practice (Mogull 2018). For healthcare authorities, the ability to communicate effectively is, and will be ever more important to deliver the most informed and (by extension) beneficial outcomes to their patients.

The General Medical Services (GMS) contract and the Quality and Outcomes Framework (QOF) were introduced in 2004, supporting the role of general practice in public health. The GMS contract included the formation of Local Enhanced Services (LESSs), which have been particularly effective in involving GPs in locally driven public health efforts supporting a wide range of evidence-based public health activities (Peckham et al., 2015). Funded by and working in collaboration with Public Health Dorset (a local division of Public Health England), this study set out to identify how various communication interventions can influence GP referral rates. It explores the relationship between public health bodies and front line general practitioners (GPs). It looks at how organisations like Public Health England attempt to improve patient healthcare through GP practice via referral-based programs, and the communications supporting such initiatives. As a profession, the role of GPs spans both scientific-medical and public communication. As will be discussed, these two distinct domains demonstrate both contrasts and similarities in approaches to communication.

The research questions guiding this study are:

RQ1: In the context of Public Health Dorset, what (if any) effect can be identified as a result of previous communication interventions on GP referral rates to public health programmes?

RQ2: How effective was Public Health Dorset's strategy in improving referral rates among general practitioners?

1.1 Motivation

In 2015, the UK Government advised local councils to pursue programmes that would improve the sustainability of local healthcare efforts for years to come. In this capacity, council funding is applied to public health authorities, which concern themselves with the incidence, distribution, and possible control of diseases and other factors relating to health. Baum (2016) describes the impact of public health as "relevant to all countries, developing, transitional, or industrialized." Improvements in efficacy in this area can, therefore, be considered a vehicle to offer far-reaching benefits to society.

Throughout the UK, regional, public health organisations funded by local governments are tasked with the continued adaptation and evolution of population wellbeing. One such authority addressing these epidemiological concerns is Public Health Dorset (PHD). Throughout its namesake county, the organisation refers to their operations as "Working as part of Bournemouth Borough Council, Dorset County Council and the Borough of Poole, we want to achieve Prevention at Scale – helping as many people as possible to stay healthier for longer; which is a key aspiration of Dorset's Sustainability and Transformation Plan (STP)." This broadly defined activity is referred to as Prevention at Scale, where three specific service gaps are identified (Dorset Clinical Commissioning Group 2016);

- The health and well-being gap
- Care and quality gap
- Finance and affordability gap

Much of this service is in an effort to work toward sustainable improvement of public health within the organisation's geographic authority.

Activities already undertaken by PHD that are considered relevant to this study include the gathering, processing and analysis of data from various healthcare institutions, trusts and public health programmes. Example outputs of these activities are understood to be used to forecast needs, inform decisions and work toward the goals set out in the counties' response to NHS England: Sustainability and Transformation Plan (2014).

The ultimate motivation being the enablement of such organisations to better utilise information they already have to provoke action by the target group (GPs) toward improved public health - a desired outcome where the means to achieve it is not always communicated effectively.

1.2 The Problem

The proposition of a national health system represents the modern face of UK healthcare, since its establishment the service has been evolving for over 70 years (Greengross 1999) into the complex primary healthcare service seen today. Over this period, motivated in part by the continued development of medical practice, GPs exist in a community of medical professionals, which, as a group, has become more specialist (Sur 2011) through the application of increasingly sophisticated medical knowledge and evidence-based research. The development of this knowledge has a less clear relationship to public health, and various crossovers of actions which rely heavily upon successful communication practices become apparent.

In the UK, an aging population, decreases in the number of full-time equivalent GPs per 100,000 patients, and an increase in per-patient consultations has resulted in growing pressure on primary healthcare (Hobbs, 2016). Riley (2018) studied stress factors affecting GPs, finding that increasing demands placed upon their service delivery, their inter-professional relationships and expectations placed on operational procedure were all highlighted as major sources. Time with patients is one aspect of these pressures, a GP-patient interaction is on average 9 minutes 2 seconds - as multimorbidity and an aging population increase demand, patients want more time, when GPs have less available (Irving 2017).

Continued professional development (CPD) places further demands upon GP time. This process represents an essential part of maintaining relevant and effective healthcare. However, the body of knowledge from which general practice is built upon is ever expanding, and with increasing complexity (Sur 2011). Despite having access to many types of furthering CPD, GPs are often untrained as health educators, and may have a narrow view of health promotion and limited experience of community development activities. Interviews conducted with several GPs by Peckham et al, (2015) revealed a lack of knowledge about wider health improvement issues and services that were not practice-based. GPs' knowledge about other services in the local community that provide health improvement were found to be variable. For GPs to keep abreast of what is already an expansive field of knowledge must be factored into their working practice and interactions. Engaging primary care and keeping practitioners appropriately informed requires tailored approaches (Peckham et al., 2015)

The complexity of healthcare communication (Mogull 2018) also contributes it's own challenges in the dissemination of information by public health services, via GPs or dedicated public health authorities. This aspect and those supporting it are broadly the areas of investigation of this research project. In a review of health promotion opportunities for general practice, Watsons, (2008) noted that 'there has been a dearth of information about the effectiveness of health promotion in the primary care setting'.

1.3 Scope

The question of developing a strategy toward better GP communication from sources outside the target group is wide-reaching. There were aspects of both the research project and the focus organisation (Public Health Dorset) that acted as logical limiters to how expansive this study would be.

Considering human communication as a ubiquitous phenomenon, the study looks at the effectiveness of professional medical communication, the attributes of what is unique to this modality of communication and how they can be monitored in a quantifiable manner for the measurement of change. This is considered in the context of a cross-organisational, interprofessional environment, between public health and primary care practitioners. This method includes categorising key factors that are influenced by scientific underpinning, such as numerical notations and data presentation, through to sentence formation and complexity of language use.

Note that the study does not attempt to propose new methods of professional medical communication, drawing only on existing research to measure and contextualise effectiveness. It therefore addresses how these known attributes can be measured with the intent to optimise communication practice and affect operational change with GP's.

The study was conducted over a nine month research period, in this time all data acquisition, case study organisation investigation and methodological research had to occur. The availability of key decision-makers and staff who were directly involved with the development of the programme, and their ability to provide archival data and historiographic information were critical to such a process.

Working in partnership with Public Health Dorset, the scope of the project was motivated by their desire to identify strategies that could be derived from the results of this study to enhance their operational practice. This factor did not lead the research, but it did influence some of the wider scope decisions that might not be seen in a purely academic investigation.

The geographical administration of the subject organisation would limit data to individuals within the county of Dorset. Unlike a national audience source, this point represented an immediate limitation of the populous characterised in the archival data collection. The research conducted can instead be considered a pilot study, with principals and methodological decisions that would inform a larger scale equivalent.

The participants of the study, those who existed in the archival data sources would be representative of genuine respondents. It was hoped this would offer credibility to any conclusions drawn from the research and would identify any important limitations for any subsequent studies of a similar format.

As part of the case study organisational remit (discussed previously) multiple programmes were/are in operation to achieve long-term public health goals. Focusing on one such programme allowed for a logical reduction in the scale at which this study would need to concern itself. One such programme which most suitably reflected the problem space was titled Live Well Dorset (LWD). Launched in April 2015, as part of a restructure, various well-being services were unified by Public Health Dorset under this one collective name. The Live Well Dorset service is freely accessible to the public who can register through various channels - one of which is in consultation with their GP. Through informal meetings with Public Health Dorset staff, GP based referrals had been highlighted as a seemingly underperforming source (in reference to no-specific expectation), and therefore offered a desirable target for the study. The programme exists as a combined service, encompassing 4 pathways; smoking cessation, weight loss, alcohol intake reduction and increased physical activity. An additional point of note at the time was the development of a bespoke case management system that included an innovative management process designed around the COM-B behaviour change model (Michie 2014). This aspect was considered a unique advantage of the service to both practitioners and patients.

Proposed by Michie (2014), COM-B is used for the planning of practical behaviour change interventions. The model proposes that an understanding of an individual's 'Capability', 'Opportunity' and 'Motivation' should be used to direct the appropriate 'Behavioural' change intervention. In the 7 years since its proposition, it has gained popularity, but measurable implementations were limited at best.

The Behaviour Change Wheel: A Guide to Designing Interventions (Michie 2014) followed the original proposition of the framework as a guide to its practical application and implementation. This resource was used by Live Well Dorset during the development of their client management system at the centre of this study. In the five years since publication, few other practical examples were found that offered this level of quantifiably tracked usage across multiple well-being pathways.

In 2018 Newcastle University's Centre for Translational Research in Public Health completed a study into the effectiveness by which the COM-B model had been implemented for the Live Well Dorset programme. The mapping of the theory was deemed to be to an excellent standard, however, the report was specific in its purpose and did not look to review the impact of the services it applied to (Rodrigues 2018).

The COM-B aspect of the programme offered a relatively unique opportunity in the direction with which to take the study. Initial research was conducted to understand the potential impact of this aspect as part of the wider research question and formed part of the literature review but was later discontinued due to the relative lack of credible data.

Public Health Dorset regularly works with various healthcare data sets collected across their local authority, these are managed and presented via the web-based business intelligence service, Tableau. Initial investigation showed the data collected by the LWD system had already been presented in an analytical form to the extent of top-level categorisations and totals with regard to the various pathways of interest. Tracked quarterly, these included a number of pathway enrolments such as client progress against arbitrary goals such as weight loss (>5%) or a reduction in cigarettes smoked, etc. There was therefore extensive scope for a more rigorous understanding of correlations between client demographics, outcomes, clients who go onto adopt other services and the aforementioned GP referrals.

Attributes of both the LWD programme and cross-organisational communication by which it was delivered introduced analytical complexity. No systematic process to relate these sources of information was apparent from the discussions that took place. It was also of note that the organisation desired that any research to ideally achieve a practical outcome which could be applied to improve the service through procedural change, should potential benefits be identified.

In consultation with key decision makers within the case study organisation, it was identified that no formal working practices were in place to standardise the professional communication efforts undertaken by programme organisers. More specifically, programme managers would agree on an information delivery method and undertake the steps needed to deliver it without a method of measuring effectiveness. This activity was exemplary of the need for this research project and how it may improve the process toward better public health service delivery.

1.4 Outline

Chapter two concerns itself with a review of existing literature, this begins with specific identification of practical attributes that contribute to, detract from, or are incomplete when identifying effective professional medical communication. The literature review identifies a lack of domain specific research for measuring the effectiveness of communication activities across large heterogeneous structures. It subsequently looks to other domains of knowledge such as UK healthcare as highlighted earlier in this introduction.

The considerations and decisions made during the development of the research methodology are detailed in chapter three. The focus organisation, Public Health Dorset, and their range of sources made a case study strategy applicable, this was seen to utilise the data sets available for statistical modelling. To complement this, the explanatory nature of the archival material available was undertaken in an inductive approach. Three core areas were identified as relevant to the research and analysis; attributes of professional medical communication, chronology of communication interventions, referral counts and demographics.

The fourth chapter systematically defines and details the results of the research. These are divided in a similar structure to the methodological approach considerations; the captured historical communication activities as an intervention chronology, followed by referral stratification for demographic sanity checks and time series analysis in the form of trends and seasonality.

Chapter five discusses the results, their relative significance, these results regarding/of the referral rates by GPs are reviewed in context of their relationship to communication provided by public health Dorset. The discussion looks at data relating to the impact of several variables of communication such as type, and timeframe, on these outcomes. Additional anomalous findings obtained while running the aforementioned analysis that are felt to be of potential relevance or interest in the context or future direction of this study, are also reviewed.

Chapter six re-visits the results and discussion in such a way as to draw conclusions when evaluating the question this study set out to answer. The overriding feature of this research highlighted the importance of quality of communication tracking at an organisational level, and this point undermines the ability to accurately measure effect. To this end, further research is proposed.

2. Literature Review

2.1 Introduction

This chapter identifies attributes of professional medical communication. It includes the impact on practice outcomes and the perception of the practitioners with which they are associated (a more commonly understood paradigm).

The result of a nonsystematic review chapter also attempts to establish the factors affecting professional medical communication in an inter-organisational context. It includes communication practices shown to invoke behavioural change amongst domain specialists. Where possible, this literature was approached with a particular focus on general practitioners within the UK healthcare system, as described in Chapter 1.

The chapter subsequently reviews possible ways to facilitate the propagation of new ideas and procedures through continued evaluation of efficacy. To this end, the chapter identifies a body of knowledge focused on the practical application of communication in interprofessional healthcare settings. Identification and measurement of communication interventions and the effect they can have in healthcare settings as identified in RQ1 from the previous chapter. These factors are to be understood in the context of complex organisational structures and the multiple systemic interactions discussed in Chapter 1.

It is important to consider that within scientific and medical communication, various levels of expertise and comprehension within specialisms exist. Effective communication caters to this diversity by considering the audiences' level of understanding of both specific subject matter and language (Mogull 2018).

Spiegelhalter's (2019) publication exemplifies many of the features of effective communication discussed later in this literature review - the use of accessible language and explanation of complex topics without the loss of technical credibility. These are backed up by clear summaries to elucidate key points and the use of commonly understood phenomena as examples.

2.2 Professional Medical Communication

2.2.1 Background

Professional medical communication literature presents a large corpus of interrelated information, the overview which follows, represents a cross-section of the factors that can be drawn upon to improve communication with general practitioners within this target domain. As a proxy between public health providers and the public they serve, the review considers the importance of interprofessional communication. This topic is explored in a way that informs the problem in question - what attributes of communication can influence GPs, specifically in

decisions surrounding the referral of patients to public health services.

Factors which affect GP referral rates will be heterogeneous by nature, intuitively, even a basic understanding of the role in which they play in healthcare, and the healthcare organisations within which they operate can be considered an example of a complex adaptive system (Goodwin 2013). General practice within the UK takes various organisational forms where the consistency of internal and external communication can differ from practice to practice and GP to GP. Further influences on these practices include varying staff, underlying processes that support these services, and on a larger scale, the range of professional bodies which govern and promote healthcare practice and their relationship to public health (The Kings Fund 2017). In practical terms, the systems described result in polymorphous communication considerations, forming complex linkages between information publishers and their diverse targeted recipients (Figure 1).

Although wider influences, such as national policy frameworks, seem to influence physician practice to some degree, encouraging more focus on preventative health, the actual evidence is limited (Rice & Smith, 2002). It is difficult to truly evaluate the effect of a policy intervention on an outcome, such as GP referral rates, due to so many confounding factors that influence it. However, in their study on organisation and delivery of health improvement in general practice and primary care, Peckham et al (2015) noted several complexities to the relationship between general practice and public health. Through their interviews with UK GPs, Peckham et al (2015) found that there were varied opinions on whether their role should prioritise clinical intervention as secondary prevention, or focus on health improvement and primary prevention.

GP motivation for partaking in any public health initiatives varies widely (Peckham et al, 2015). Similarly supporting this notion, in a Europe-wide survey of GPs in other countries, the lack of reimbursement was a key reason for not engaging in health promotion activities (Brotons et al, 2005).

Though these concerns are not the focus of this study, they are mentioned here as they are felt to highlight the challenges faced when assessing an intervention at this scale.

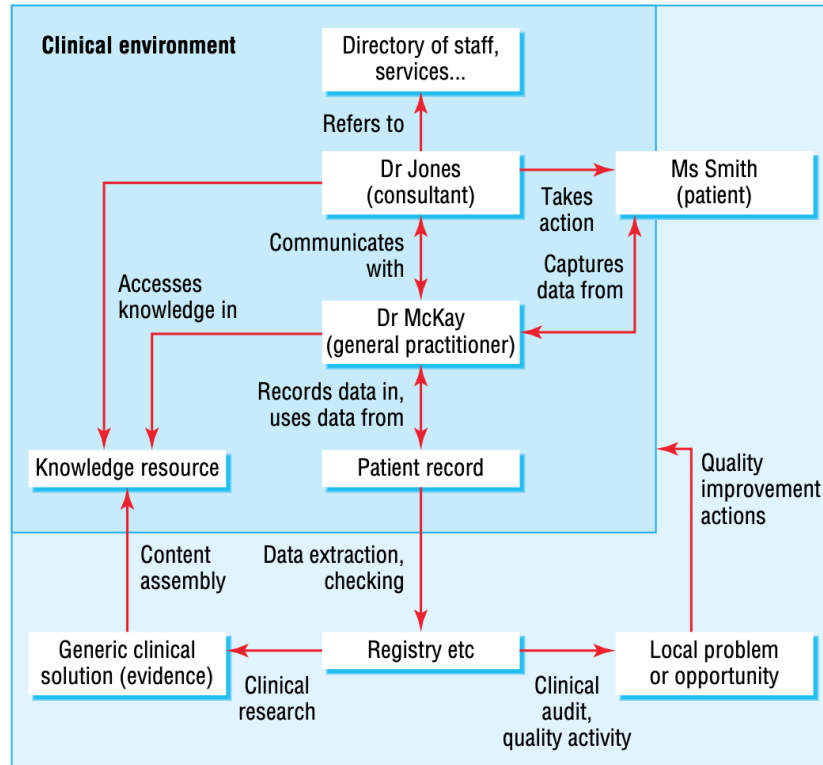


Figure 1. Wyatt and Sullivan (2005) “Information flows in clinical and non-clinical environments.”
democrats the complex interage interfomation can experience within clinical environments.

The volume of literature sees a disproportionately small number of studies conducted in the field of professional primary care communications compared to those that address public to primary care interactions. This disparity can be explained by considering the relative audience sizes of primary care physicians to their patients - in the UK for example, the NHS employs 150,273 doctors and 40,584 general practitioners (GPs) who service a population of 54.3 million people (NHS 2016).

Peckham et al. (2015) suggest GP activity is, at least partially, driven by specific contractual incentives and conditions. However, factors such as peer pressure, relationships with public health departments, education and training also influences practices. The same study (Peckham et al, 2015) interviewed GPs who expressed concern regarding the large number of changes to the overall structure and control of public health services over the past two decades. They felt that these changes risked further fragmentation of relationships between various organisational groups, as well as a lack of clarity of their role within the public health realm, leading to a sense of distrust.

2.2.2 Attributes and outcomes

The combination of negative outcomes identified by Vermeir (2015) and the key elements of professional medical communication initiatives proposed by Schiavo (2007) offers a generalised

overview for consideration when developing professional healthcare communication materials (Figure 2).

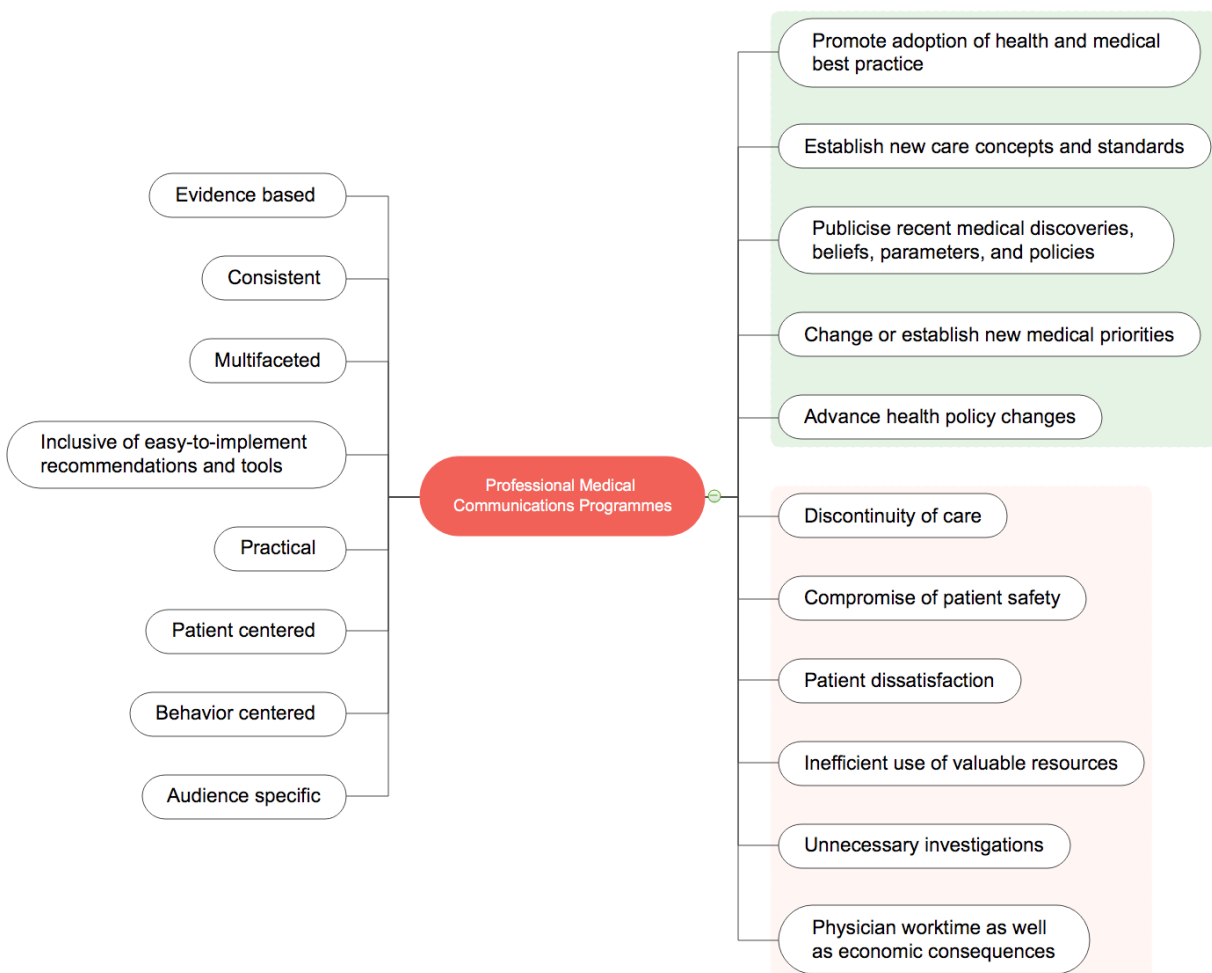


Figure 2. An illustrative association of input attributes of professional medical communication (left) as described by Vermeir (2015), and example outcomes (right), both positive (green) and negative (red), as identified by Schiavo (2007).

When considering the outcomes (Figure 2) in a context of motivational factors, there is the opportunity to frame healthcare programmes such that they will deliver one or more of the positive outcomes, or prevent one or more negative outcomes. In other words, despite the best intentions, poor communication practice is not only suboptimal in achieving its intended outcome, in some cases, it can also have wider negative impacts to healthcare practice. Opportunities for improvement, therefore, could have a significant impact on the improvement to healthcare.

Medicine and medical practice as a domain of knowledge has developed over the past 25 years

as a direct result of empirical research in the form of evidence based medicine (Sur 2011). Results produced through scientific evaluation have underpinned modern practice and process - the language of science is, by extension, also the language of medicine. This aspect separates medical professionals from their patients (the public), while at the same time binds them inter-professionally through a common lexicon and discourse. Discussion of the merits and shortcomings of evidence-based medicine are outside the focus of this document and should be considered an expectation within the medical domain.

2.2.3 Inter-organisational communication

Communication in professional healthcare settings often relays complex data and statistical information that may be presented in many formats. Different interpretations of this information has the potential to influence the comprehension, and even the meaning conveyed by the figures. In this sense, consideration is given to domain-specific conventions. Akl (2011) identified that health risk outcomes were “better understood when [...] presented as a natural frequency rather than a percentage. e.g. ‘65 of our 215 respondents’ rather than ‘30%’.” It was also concluded that “on average, people perceive risk reductions to be larger and are more persuaded to adopt a health intervention when its effect is presented in relative terms” e.g. ‘one third better off when you do X’. The influence of statistical format revealed no discernible difference between professionals and consumers.

Looking at existing authoritative communication guidance for the current state of recommended professional medical communication practice, there are numerous examples of clinical commissioning group (CCG) strategies and guidelines. They focus primarily on patient communication and their content is not anchored in academic literature or research. A review of the volume and variation available across each healthcare authority by which they are written could represent a systematic content analysis research project of it's own. Of the documents reviewed, the general understanding of challenges faced in healthcare communication are highlighted, but in aggregate appear as though they offer relatively nonspecific strategies rather than practical, actionable advice. This situation poses questions around the quality and consistency of advice being made available. This point itself is a large focus of this research. “Tools and routes” represented an example that includes detail on interprofessional communication, and referenced survey results to justify the claims, but is still not formulaic in nature.

Vermeir (2015) conducted a systematic review which identified face-to-face communication as being the most effective at producing the desired result - behavioural change. In part, this was due to the additional information body language conveys and the immediate ability to clarify complexities. There are obvious limitations to this means of communication making it regularly infeasible, especially at scale. Although, geography has become less of a concern as real-time video service becomes more readily available. A “consensus on particular advantages of written

communication over face-to-face communication” was also identified in the study, namely, immediacy (in digital form) having no requirement for arranging meeting times - as would be the case with face-to-face communication. Reproducibility was also seen as an advantage and, by extension, the ability to reference written assets. There are further advantages dependent on the method of distribution, such as email, where the ability to track engagement represents a further advantage.

In the context of modern society, email is commonly accepted as a largely ubiquitous channel of information delivery, both inside and outside the medical service domain. As such, it is also a desirable choice for professional healthcare communication. However, despite being such a widely used channel, its impact relative to alternative modalities is not properly understood. Goyder (2015) found no evidence base to derive meaning from the use of email between healthcare professionals, suggesting rigorous studies would be required in this area. Berendsen (2009) study of the communication between 259 GPs and 232 specialists showed the use of multiple communication mediums including email. Disparities in the perception of effectiveness between the groups suggests an element of ‘better than average’ bias exists and the value of self-evaluation responses to questionnaires should be approached with caution. Both GPs and Specialists showed 94.9% and 89.0% respectively that feedback was an important factor. This finding is a particularly prominent feature of the study and points to considerations that there is more to the process of effective communication than the point of intervention.

Mogull (2018) describes effective communication as presenting “a clear, logical, and persuasive argument articulating a case in which the evidence information (data) justifies a particular insight, interpretation, or conclusion.” The review of over 500 published research findings is itself exemplary of the complexity of scientific medical communication (SMC). Information in this form would not be feasibly transferable to health staff without them becoming specialists in research publication.

The way in which evidential authority is derived - through research and studies that employ scientific method - requires demonstrable statistical significance in the findings. This process, at scale, guides decisions that become national healthcare policy. Scientists who write and subsequently publish this research are rarely specialists in communication in addition to their field of expertise. They typically learn both writing style and language from the professional materials they have consumed previously through education and research. These materials are themselves written and published by other scientists, rather than persons specifically trained in communication. A process of imitation follows, creating a feedback loop of communication practice. This aspect makes scientific and medical communication susceptible to the self-perpetuation of poor practices and increasingly inaccessible language. In other words, many publications are not optimally designed and written for their primary purpose - transferring knowledge (Mogull 2018).

A tangible example of the differentiation between specialist medical communication and public discourse can be identified by the Lexile score comparison of popular media (newspapers) and

that of scientific journals diverge over time (Knight 2003 cited by Mogull 2018). The Lexile score is used to measure reading comprehension; when evaluating text, a score is given based upon algorithms which “analyzes the text’s semantic (word frequency) and syntactic (sentence length) characteristics and assigns it a Lexile measure” (MetaMetrics 2019). In the case of Knight 2003 study, This divergence in language complexity (Figure 3) shows, in part, how scientific communication has become continually less accessible to the public over the last century. There is, however, no indication of the rate at which popular language has transformed against itself over time, that is to say it will change independently in its own right. The simplification of language can be used to make content indiscriminately accessible to the broadest populous (Agarwal 2013). In that sense, the rate at which divergence has increased could be partially a result of *dumbing down* in mass-media as an attempt to expand audiences, increasing the gap between popular material and scientific literature.

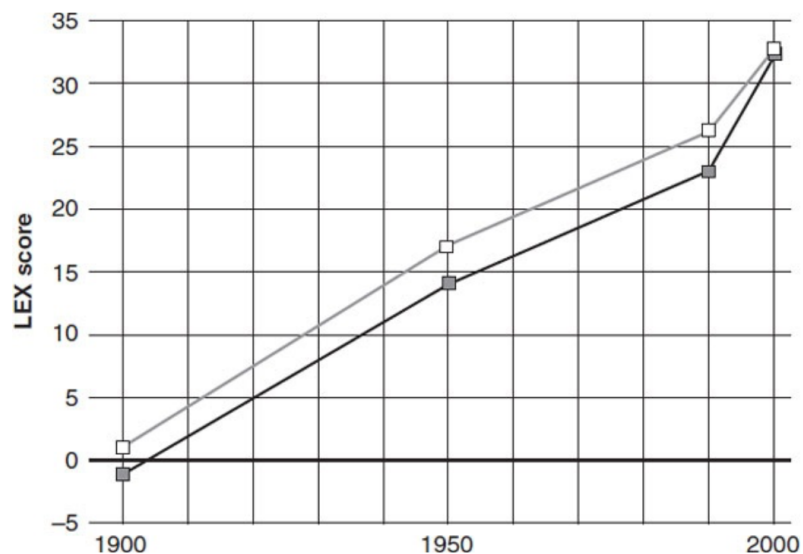


Figure 3. The lexile scores (used to measure the difficulty level of reading a text) of the publication *Science* (top line) and *Nature* (bottom line) are seen to have dramatically increased compared to the baseline (0) of popular language defined as that used by newspapers over time. Source: Knight (2003).

“For healthcare providers, poor communication leads to additional workload as it decreases confidence in decisions” (Vermeir 2015). While this statement is in the context of referrals, it would be logical to suggest that additional workload is always seen as a negative if possibly avoided. It is, therefore, inferred that the opposite outcome can be considered a positive, and can be a target of effective messaging. In reference to message framing, it would suggest that positive effects (specifically around reduced workload) could offer the incentive needed to improve uptake of public health communication.

Information can be framed both positively and negatively in its own right. People tend toward risk aversion when presented with positively framed information, and tend to seek risk when presented with negatively-framed information (Tversky 1981 cited by Akl 2011a). The results of Akl (2011a) systematic review did not, however, identify a consistent impact on outcomes in either case. This finding suggests that both approaches are as effective as one another, or as ineffective, depending on the observer's frame of reference.

Within a practice-based environment, GPs typically work alongside various other medical professionals; nurses, practice managers, other GPs. Dadich and Hosseinzade (2016) examined the channels by which “primary care clinicians learnt of resources on evidence-based sexual healthcare.” The study found educational events and colleagues to be the most significant source of awareness. However, the method by which clinicians learned of resources did not appear to influence the perceived impact or, by extension, use of the resources in question.

Cognitive speech actions represent a specific class of interaction, “expressing an expert’s opinion” which serves to “impart information and as a vehicle for the production of new professional knowledge among peers” (Akl 2011b). The review also clarifies that “It is conventionally accepted to soften categorical judgments and negation to develop communicative process in an effective way.” Professional conversations between GPs in informative sessions have a formulaic structure and a basic narrative construct (Shamne and Nevzorova 2017; Horder et al, 1986). Polite and well-structured language is typically expected, slang and other short forms should therefore be avoided unless contextually significant.

Looking toward other domains that have helped progress data communication best practice, specific similarities can be seen. Knafllic (2015) describes data communication through narrative structures for increased engagement. Aspects such as treating the viewer as a contemporary as having similar/shared knowledge should not be assumed, echoing that of variation in specialist knowledge as mentioned previously (Mogull 2018). Interestingly, there are noticeable stylistic differences of the data communication literature itself, being non-medically oriented, vs that of SMC literature. It is typically less formal, displaying more accessible language - in a self-referential way, validating the aforementioned concerns of complexity found in professional medical communication.

The ideas around vocatives, appellatives and wider formal communication structure point, again, toward the importance of professional peer backing that has previously been highlighted as a key concept by Dadich and Hosseinzadeh (2016) - “The perceived opinions of peers and opinion leaders play a major part in influencing the attitudes of individual practitioners and, most importantly, their decisions to act on new information.” Historically, widespread changes to GP practice take time, as a general consensus first forms between peers and subsequently the wider domain-community of practitioners. Longitudinal data would need to be available to analyse effects of this type. Identifying quantifiable associations between both multimedia and multimodal communication materials. Discernable shifts in referral behaviour represents a

means to developing the required understanding of how interprofessional medical communication attributes might best influence GP referral practice.

Categorisation	Approach	Explanation	Example/Action	Source
Strategy & Structure	Communication basics	Medical communication may be a specialist domain, but the target audience are still human, more generally proven communication techniques are still relevant.	Continue to use well know and established communication techniques as long as they do not counteract the SMC specific points.	Mogull, S.A., 2018. Scientific and Medical Communication: A Guide for Effective Practice (ATTW Series in Technical and Professional Communication). Akl, E.A. et al., 2011. Using alternative statistical formats for presenting risks and risk reductions Cochrane Consumers and Communication Group, ed. Cochrane Database of Systematic Reviews, 58(8907), pp.455-90.
Strategy & Structure	Begin with context	Be clear from the outset exactly what you wish the audience to know and the data available to support it. Irrelevant or elaborative information may introduce misinterpretation and detract from the purpose of the material.	Before producing materials, spend time defining explicitly why you are producing them.	Knaflic, C.N., 2015. <i>Storytelling with data: A data visualization guide for business professionals</i> . John Wiley & Sons. Rose, D., 2016. <i>Data Science: Create Teams that Ask the Right Questions and Deliver Real Value</i> . Apress.
Strategy & Structure	Be concise	Medical professional discourse typically adheres to at least simple presentation structures.	Avoid including additional information that only 'might' be useful, stick to the core message.	Mogull, S.A., 2018. Scientific and Medical Communication: A Guide for Effective Practice (ATTW Series in Technical and Professional Communication).
Strategy & Structure	Minimum formal structure	A typical story structure will connect and guide the audience through the content to an actionable conclusion.	Minimum; introduction, main part and conclusion.	Horder, J., Bosanquet, N. & Stocking, B., 1986. <i>Ways of influencing the behaviour of general practitioners</i> . J R Coll Gen Pract, 36(292), p.517.
Strategy & Structure	Narrative structure	Identify and explain the specific action that should be taken to achieve the messages desired action.	Consider introducing the messaging develop characters and events culminating in a desired action.	Rose, D., 2016. <i>Data Science: Create Teams that Ask the Right Questions and Deliver Real Value</i> . Apress.
Strategy & Structure	Specify actions	Uptake of practice change is more difficult when actions fall outside the audiences existing values.	Include clear, simple and practical instructions.	Grol, R. et al., 1998. <i>Attributes of clinical guidelines that influence use of guidelines in general practice: observational study</i> . BMJ, pp.1-4.
Strategy & Structure	Avoid controversy	There is no one better placed to advise upon domain specific content elements than the audience themselves.	Gain understanding of the target audience with reference to the proposed messaging action.	Grol, R. et al., 1998. <i>Attributes of clinical guidelines that influence use of guidelines in general practice: observational study</i> . BMJ, pp.1-4.
Strategy & Structure	Audience review	Communication is evolutionary in nature as technology and accepted norms change over time.	Have comms materials reviewed by example target audience individual before distribution	N/A
Strategy & Structure	Evolution	Formal language offers familiarity and confidence, but it should not be unnecessarily complex because this can introduce barriers to cross-discipline propagation.	Track, measure and review communication	Mogull, S.A., 2018. Scientific and Medical Communication: A Guide for Effective Practice (ATTW Series in Technical and Professional Communication).
Text	Formal yet accessible language	Expert opinions as recommendations rather than obligations are considered more polite and will be more familiar, therefore building trust.	Understand and include formal language, but avoiding letting this introduce unnecessary or complex prose. "X suggest that...", "X proposes..." etc.	Vermeir, P. et al., 2015. <i>Communication in healthcare: a narrative review of the literature and practical recommendations</i> . International Journal of Clinical Practice, 69(11), pp.1257-1267.
Text	Recommend expert opinions	To improve familiarity and confidence, respect the variation in confidence and understanding when offering expert opinions.	"It is believed that...", "It is understood that..." etc.	Shamne, N.L. & Nevzorova, M.S., 2017. <i>Lingua-pragmatic potential of speech actions in medical professional discussions</i> . In 7th International Scientific and Practical Conference "Current issues of linguistics and didactics: The interdisciplinary approach in humanities" (CILDIAH 2017). Paris, France: Atlantis Press, pp. 1-8.
Text	Soften categorical judgments	Domain specific language can promote familiarity and confidence, but should be limited to generally understood topics - very specialist knowledge may not be widely understood.	Use common technical terms/acronyms/phrases where suitable. Reviewed by a domain specialist when possible.	Shamne, N.L. & Nevzorova, M.S., 2017. <i>Lingua-pragmatic potential of speech actions in medical professional discussions</i> . In 7th International Scientific and Practical Conference "Current issues of linguistics and didactics: The interdisciplinary approach in humanities" (CILDIAH 2017). Paris, France: Atlantis Press, pp. 1-8.
Text	Include technical language	Support credibility of information using references and evidence from organisations which GPs consider to be authoritative.	If the evidence is from an authority, reference them. If not, find an alternate that may be able to back the evidence.	Dadich, A. & Hosseinzadeh, H., 2016. <i>Communication channels to promote evidence-based practice: a survey of primary care clinicians to determine perceived effects</i> . Health Research Policy and Systems, pp.1-12.
Text	Reference authority	Positive and negative goal framing, use language to described the outcome of doing or not doing an intended action and the subsequent positive or negative outcomes.	"...patients health improved over a six month period." or "...patients who did not receive treatment X continued to deteriorate."	Akl, E.A. et al., 2011. <i>Framing of health information messages Cochrane Consumers and Communication Group</i> , ed. Cochrane Database of Systematic Reviews, 22(1), pp.60-84.
Text	Framing	Where possible, avoid unnecessary complexity, not all numbers require graphs.	With only a number or two, highlight, use clear prominent text instead of resorting to graphs.	Knaflic, C.N., 2015. <i>Storytelling with data: A data visualization guide for business professionals</i> . John Wiley & Sons.
Visuals & Graphs	Numbers	Use color, size, and position to visually prioritise content.	Use statements like "86 out of the 1142 patients studied..." rather than "7.5% of the patients studied..."	Knaflic, C.N., 2015. <i>Storytelling with data: A data visualization guide for business professionals</i> . John Wiley & Sons.
Visuals & Graphs	Hierarchy	Use natural frequencies rather than percentages when both options are valid.	Express 12 out of 20 as 60% (being a larger more impactful number).	Akl, E.A. et al., 2011. <i>Using alternative statistical formats for presenting risks and risk reductions Cochrane Consumers and Communication Group</i> , ed. Cochrane Database of Systematic Reviews, 58(8907), pp.455-90.
Numbers	Natural frequencies	Bigger numbers of the relative changes have more impact than smaller absolute changes for the same results.		Akl, E.A. et al., 2011. <i>Using alternative statistical formats for presenting risks and risk reductions Cochrane Consumers and Communication Group</i> , ed. Cochrane Database of Systematic Reviews, 58(8907), pp.455-90.
Numbers	Relative change	Absolute values are more suitable to express risk by expressing both the magnitude of reduction and the baseline risk.		Akl, E.A. et al., 2011. <i>Using alternative statistical formats for presenting risks and risk reductions Cochrane Consumers and Communication Group</i> , ed. Cochrane Database of Systematic Reviews, 58(8907), pp.455-90.
Numbers	Values of risk	Empirical evidence that shows meaningful trends can inspire action.	Identify and draw upon evidence which is worthy of attention, avoid evidence that does not carry impact.	Schiavo, R., 2007. <i>Health communication: From theory to practice</i> . John Wiley & Sons.
Numbers	Statistically significance	Using the correct domain specific notation on evidence will imply legitimacy and improve confidence.	mg = milligrams	Schiavo, R., 2007. <i>Health communication: From theory to practice</i> . John Wiley & Sons.
Numbers	Scientific notation	Without clear explanation empirical evidence will not be considered credible and can in turn detract from confidence.	Include clear source and relevance information when necessary, if this is not possible find alternative evidence.	
Numbers	Clear labeling	Offering financial incentives can motivate some audience, however an already oversubscribed GP will be less incentivised to add more to their workload.	Review opportunity to incentivise messaging action through financial reward and implement if deemed to be contextually suitable.	Horder, J., Bosanquet, N. & Stocking, B., 1986. <i>Ways of influencing the behaviour of general practitioners</i> . J R Coll Gen Pract, 36(292), p.517.
Motivation	Financial incentives	GP time is stretched and therefore a valuable commodity, reduced workload can be a strong motivator.	Incentivise intended action in the messaging (that can result in saved time) by identifying this as a possible outcome.	
Motivation	Time saving	Endorsment from colleagues or peers offers greater credibility.	Consider appealing to secondary audiences, make materials shareable.	Horder, J., Bosanquet, N. & Stocking, B., 1986. <i>Ways of influencing the behaviour of general practitioners</i> . J R Coll Gen Pract, 36(292), p.517.
Delivery	Colleagues & Peers	Multimedia delivery and commonly used by GPs, fast, shareable and trackable.	Develop email content to send to multiple recipients, use suitable campaign services to allow taggin for tracking of specific campaigns open rates and click through rates	Vermeir, P. et al., 2015. <i>Communication in healthcare: a narrative review of the literature and practical recommendations</i> . International Journal of Clinical Practice, 69(11), pp.1257-1267.
Delivery	Email distribution	Continued personal development is a common part of GP practice, this environment is focused and credible.	Include messaging and	Horder, J., Bosanquet, N. & Stocking, B., 1986. <i>Ways of influencing the behaviour of general practitioners</i> . J R Coll Gen Pract, 36(292), p.517.
Delivery	Educational events			

Figure 4. Matrix of communication methods and actions in a primary healthcare setting gathered from resources.

2.2.4 Summary

Across the various sources of literature identified in this chapter, there was seen to be a lack of evidence of the influence on physicians toward preventive care (Rice & Smith, 2002) and generally within healthcare communication environments. A point unambiguously highlighted by Peckham et al. (2015) "There is an urgent need to develop better-quality and more relevant research studies that examine the way interventions are delivered and organised to support

continuing developments in health promotion and prevention that are being prioritised in policy and practice.”

The chapter looked at the specific attributes which can be considered practically applicable in GP communication practices. The methods of intervention identified within the literature as applicable to the research are extrapolated with explanation and transposed into a matrix (Figure 4). The matrix offers a collection of communication methods featured within the literature source studies, but there is little in the way of a unifying theory or application in the context of a medical policy environment.

3. Research Methodology

3.1 Introduction

This chapter describes the concepts, decisions and methods undertaken during the research study. The methodology is expressed in several sections, which cover strategy, approach, data collection methods and tools, and considers these in the context of existing research practices. Following this, there is a short section describing the limitations and ethical implications.

The study set out to identify how public health communication interventions can influence GP referral rates. As identified in Chapter 2, it became clear that the circumstances under which such processes occur are (in-part) representational of the factors by which referrals might be influenced. It was decided that tests undertaken in controlled conditions would offer abstract findings that may not truly reflect the efficacy of primary care communication in practice, and be less relevant outside the conditions under which they were collected.

The intent to invoke action amongst an audience through communication posed a desired outcome of increases in referral rates by general practitioners to public health services. Conceptually, it was proposed that relating attributes of professional medical communication to referral rate changes, in a real-world setting, would offer a suitable mechanism for measuring and quantifying this relationship. In turn, this approach would offer evidence to the efficacy of each intervention, with the purpose of identifying the most effective and efficient communication practices that maximize GP interaction and uptake. In the setting of this study, uptake takes the form of patient referrals. Interventions will be classified based on their apparent communication attributes; distribution mode, modality and content.

The focus of the study was Public Health Dorset's 'Live Well Dorset' programme, which registers, facilitates and monitors people and their progress through four wellbeing service pathways. As stated in the introduction, this organisation facilitates public access to health and wellbeing programmes. One key source of service users is those referred to by their GP. Improving this process was considered a desired outcome, therefore, Public Health Dorset commissioned this independent study of their own process and outcomes.

In summary, the methodology would need to address the identification and capture of historic communication activities that were undertaken and would affect a GPs decision to refer their patients, and to what extent/magnitude this had taken place.

The studies discussed in Chapter 2 identified a broad spectrum of communicative considerations, with experimental evidence for their efficacy on a per-attribute basis. In part, this was attributed to the relatively small sample sizes resulting from questionnaire based primary research (Berendsen 2009; Akl 2011; Dadich and Hosseinzadeh 2016). It can therefore be

difficult to see how these principles translate to emergent effects across complex organisational systems.

Having developed an understanding of the attributes that can inform the effective use of communication activities in medical practice, attention turns to how these attributes may be measured in an active healthcare environment. In the field of public health, and healthcare research in general, measuring effect is a common aspect of many academic studies. Concerning itself with the study of populations and the effect of a given treatment on those people, it is common practice to identify distributions of demographic information across a sample set, and quickly review cross-comparisons. Requirements such as these find stratification a common solution to summarising study data (Leyland 2016). The sample used for this study was an archival data set of over 17,000 cases tracked by the Live Well Dorset program, provided by Public Health Dorset

Where stratification becomes infeasible due to exponential growth in cross comparisons such as hierarchical data, multivariate (or multilevel) analysis (MLA) is required, Mitchell (2011) summarises the advantage of this approach as “It allows you to simultaneously assess the impact of multiple independent variables on outcome.” Furthermore, multilevel analysis also helps the often unavoidable nature of various sample sizes. Statistical analysis approached from a healthcare research perspective caters to a domain that requires the identification of response to medical treatments or interventions. Coupled with the communication literature reviewed previously where typically the approach to the measurement of efficacy was through the use of survey results, these do not offer a view of the overall effectiveness of interrelated, inter-organisational communication. A more general approach should be offered; one which could strategically inform healthcare organisations when developing their communication strategies.

Considering the time to see the perceivable change in healthcare practice as identified by Dadich and Hosseinzadeh (2016), it stands to reason that trend analysis would be best suited to longitudinal data (Nakai 2009). More specifically, in the case of GP referrals, time series data offers a suitable line of enquiry, as an already widely researched area of statistical analysis. Of the various approaches available, relating communication activity effects on GP referrals would require identification that a causal relationship could be inferred. From a statistical standpoint inferring causation is a nondeterministic consideration, as the name suggests, the decision on whether or not a causal event took place is an inference rather than fact. This approach requires comparison of what did happen and what would have happened without intervening (Spiegelhalter 2019) which can not truly be known.

Taking a step away from the healthcare domain, statistical analysis in this capacity is an area in which more commercial entities are already addressing at scale. From the literature of recent years, the ever-increasing ubiquity of digital data capture sees large media corporations, those with funding for digital services well beyond that of healthcare, investing in statistical analysis and prediction toward the identification of effects and causal inference in time series data (Taylor

and Letham 2017; Brodersen 2015). These solutions include mathematics far beyond the scope of this study, and the expertise of the researcher, however, importantly they have been actualised as usable tools (code libraries) removing the requirement for advanced statistical experience.

3.2 Research Strategy

Studies reviewed in Chapter 2 did not identify a specific method for measuring the effectiveness of primary care communication materials or programmes. It was therefore necessary to establish what constitutes effective professional medical communication. In other words, a system by which the primary care communications could be quantified.

With a focus on practical, real-world systems, while existing within the healthcare domain, Action Research was identified as a potentially suitable methodology choice. However, as the ownership and delivery of the research was to be conducted externally of the organisation, this did not represent a methodological best-fit (Denscombe 2016). Intuitively, the situation presented what is commonly considered a case study scenario, and as an approach to the research, this methodology would accommodate the variety of sources, types of data and exploratory research methods which had been identified as being necessary during the initial investigations.

The case study organisation presented various archival source data options in varying formats which were interrelated yet not analytically compatible at source. Each source would require varying levels of processing before analysis could be conducted - both quantitative and qualitative. In the wider context of the desired outcomes of the research, the methodology would also need to address the real-world and change-centric desires of the target organisation.

Structurally, the research took the form of a part-interpretation of Kothari's (2007) five phases of case study research. The interpretation and proposed implementation of each phase is introduced and described as follows:

- “(i) Recognition and determination of the status of the phenomenon to be investigated or the unit of attention.
- (ii) Collection of data, examination and history of the given phenomenon.
- (iii) Diagnosis and identification of causal factors as a basis for remedial or developmental treatment.
- (iv) Application of remedial measures i.e., treatment and therapy (this phase is often characterised as case work).
- (v) Follow-up programme to determine effectiveness of the treatment applied.”

For the purpose of this study, (i) was satisfied by the relationship between GP referral rate and communication interventions in the context of the attributes of these activities. The second

aspect of the research (ii) was proposed as the collection of archival data across three areas, which would together inform:

- Definitions of historical intervention activity (qualitative)
- Referral data (quantitative)
- Attribute classification of professional medical communication (qualitative)

The diagnosis of causal factors (iii) would reply upon analysis of the aforementioned data collected through statistical analysis where relationships between intervention events were identified. Time constraints of the study would limit the ability to achieve phases (iv) and (v), yet the results of the study were to inform such next steps for the purpose of further research.

In the first instance, a matrix of professional healthcare communication best practices was developed from the literature review corpus. To ensure relevance to the case study organisation, points extracted were considered practical in application, and applicable to Public Health Dorset's communication capacity.

The quantitative data representing the efficacy of the communications were to be achieved by identifying associative growth in the number of referrals resulting from publication within professional medical communities. These materials included attributes of effective GP communication as identified via the literature review.

To evaluate the means by which it may be possible to increase GP referral rates through communication practices was non-trivial. It was proposed that associating changes in referral trends would offer a general solution, and if these could be related to communication interventions it would provide quantifiable evidence of the efficacy of each activity. Referral trends would be identifiable through archival sources, Public Health Dorset programmes collecting information about the wellbeing service users that the programmes support.

The relative complexity of the interchange between the proposed variables and the exploratory nature of the case study resulted in no formal hypothesis being proposed, it would however be a desired outcome to identify an association between the type and delivery of inter-organisational communication and the GP referral rates. As such, an inductive approach was taken to understand what, if any, association can be identified between various attributes of professional GP communication and the resultant number of referrals made to the service(s) target by each communication activity. To develop the initial strategy, an informal meeting was conducted as part of the initial discovery process. This highlighted 3 potential sources of information, which directly influence the Live Well Dorset programme.

1. Communications aimed at highlighting the service for primary care practitioner audiences.
2. Focus group feedback from primary care practitioners regarding the Live Well Dorset system.

3. The client relationship management system that tracks client registration and progress through the programme.

In the context of this study, GPs were considered to act as the facilitators of the case study in question. They represented the interchange between communication materials and the patient referral data. Therefore, direct interviews offered an obvious approach to data gathering. However, as identified in the literature review, GPs have limited time, when also considering the extensive nature of the target audience both geographically and quantifiably each with varying schedules, it was decided that the scale of interview results required was infeasible within the confines of this study. It could be argued that a smaller sample would have been feasible, but in this particular case the decision was made to use the quantity of archival data (as per source 3 above) available over primary data. This decision presents its own limitations that are discussed in detail later in this chapter.

3.3 Research Approach

To develop a comprehension of what aspects of communication had most effect, a framework of activities, undertaken in logical order was actioned to gather enough information to meet the needs of the strategy outlined previously.

1. Identify known attributes of professional healthcare communication
2. Catalogue and classify examples of communication activities conducted by the case study organisation
3. Quantify service referrals and associated communication activities
4. Attempt to infer correlation where associations are identified

The researchers' approach to achieve these 4 activities is described in the following sections of this document.

3.3.1 Identifying attributes of professional healthcare communication

The literature review demonstrated a widely researched domain of knowledge and range of factors affecting professional medical communication that offers tangible examples of communication efficacy and augment the activities undertaken by LWD with categorical information. This would become a process of classifying common and recognised communication traits identified within the chronological interventions. To understand the interventions more formally, component attributes of communication intervention e.g., medium, use of imagery, language, structural format and any numerical references were to be extrapolated through a process of content analysis. Developing an evidence base would rely upon quantitative data which would be used to both guide and support communication activities, while also identifying the efficacy of the materials.

In practice, this approach to the categorisation of communication interventions resulted in a single cohesive matrix of applicable communication practices. Both positive and negative outcomes identified in the literature were included to allow a broader understanding of what may or may not be relevant to the growth in activities because of healthcare communication interventions.

The process of content analysis was conducted in isolation by the researcher and was not cross-validated. With additional time and resources a more stringent classification process would offer more credibility, however, the resultant artifact found in section 2.2.4 represents a valuable resource for further research.

3.3.2 Chronology of Case Study Communication Activities

It was deemed necessary to establish a chronological record of past communication interventions undertaken by the case-study organisation. The timeline would extend as long as there was data available. This would aid comprehension of the types of activities conducted and later facilitate temporal associations between interventions and variations in referral rates over the same period. The process of collecting the relevant data was to be conducted by requesting structured responses, issued via telephone and email. The information requested was formatted and reported as a digital spreadsheet matrix pre-formatted with informational criteria. Similar to a questionnaire, with open ended capacity for distinct responses. More specifically the approach prompted the contributor, by means of titled columns, to enter relevant data on a per-event basis by completing the following fields. The collection period ran for 6 weeks with periodic reminders.

- Promotion - A short definition of the communication activity
- Source/Facilitation - Who (person, organisation) that was credited with undertaking the activity
- Asset(s) - A copy of any original assets to be supplied where available
- Date issued - The date on which the activity started
- Duration - How long the activity was intended to run for
- Medium - The form of distributed
- Scope of Distribution/Audience - What was the intended scale of the activity, including demographic considerations
- Description - A more detailed explanation of the activity if relevant
- Notes - Option for the inclusion of other points of interest regarding the activity

It was considered detrimental to the analysis to discard data that was supplied in formats outside of direct entries to the matrix document. Should relevant data be returned in informal/conversational responses, in response to the request email for example, it was also to be included. In these circumstances the researcher transposed the content on behalf of the responder as best fitted the proposed collection format. Such a process was open to the introduction of possible biases, and should be considered when reviewing this data. It was

proposed that this process could be undertaken by multiple researchers and cross referenced to help identify such shortcomings but was beyond the scope of the project resource.

The time parameter represented an obvious common relationship between communication activities and the GP referrals. This acted as the interchange between otherwise heterogeneous variables - relativising one against the other. When considering communication interventions in a linear timeline, efficacy was considered to be present if the desired outcome was sufficiently perceptible through measurement and causal inference (Brodersen 2015). In the case of this study, the measurements were represented by the number of referrals by general practitioners. As such, observed changes in the rate at which referrals were registered in the LWD client system could offer evidence of causation. Subsequent findings would be extrapolated to the component attributes of communication. In aggregate these relationships would constitute a quantitative base for how public health communication practices can influence GP referral rates while also informing the suitability of measurement methodology.

3.3.3 Quantifying Service Referrals

The case-study organisation possessed a substantial archival data source in the form of electronic registration and client-tracking records for the period 1st June 2015 to 12th January 2019. The quality of data was understood to invariably be impacted by the processes by which it was collected (at source). An initial telephone meeting was conducted with a member of staff who originally guided the development of the programme and therefore possessed intimate knowledge of the systems development and usage. From the notes taken by the researcher, the data collection process was understood and developed into the visual transaction below (Figure 5).

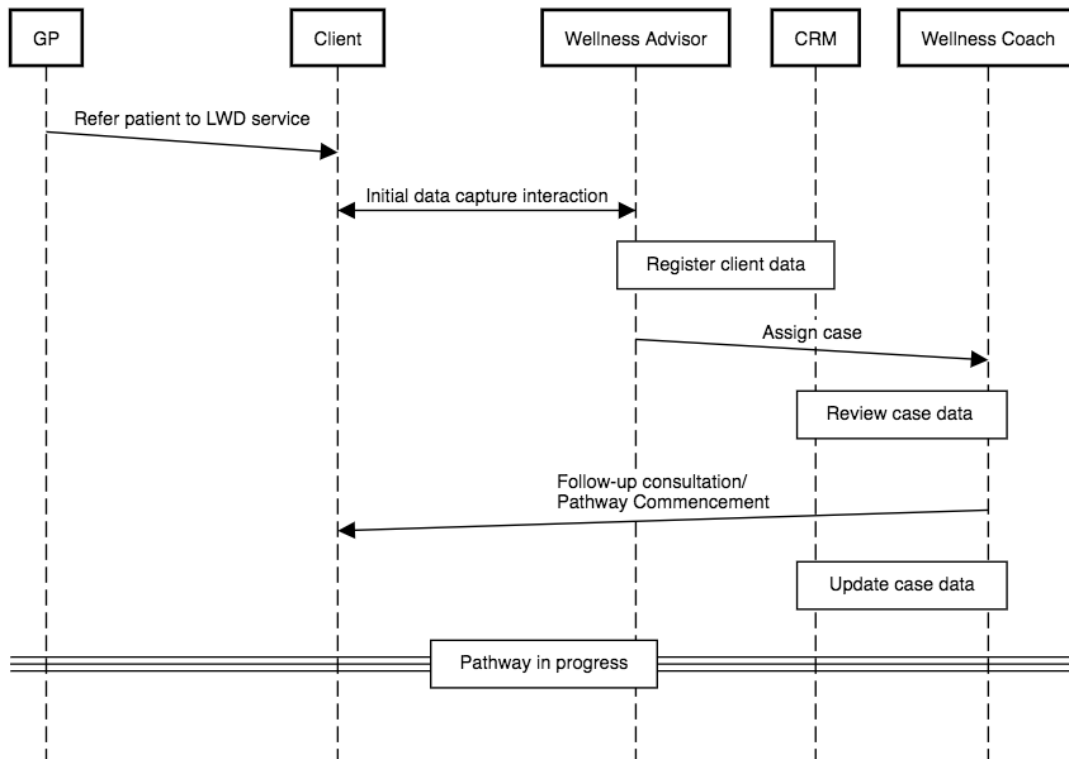


Figure 5. The sequence diagram charts the referral and signup of a client to the Live Well Dorset pathway from an initial GP referral to eventual commencement of the wellbeing pathway. Information is exchanged (arrows) and relevant interactions between the relevant parties (both humans and database), over passing time (passing from top to bottom).

The process detailed in Figure 5 begins a potential 12 month programme for participants, who are requested to check in at 3, 6 and 12 month intervals. At each check-in relevant biometrics are requested and stored within the system as a quantifiable measure of their progress.

In practice, the available data was analysed in an exploratory manner. *Fayyad (1996, p.5)* proposes that knowledge discovery in databases (KDD) is possible through a sequence of steps re-iterable as a process of identifying and refining insights not apparent in raw form (Fig 5). The approach was adopted as a methodological process and used as a de facto approach for the analysis carried out upon the supplied dataset.

Figure 6. The Knowledge Discovery Process as identified by (Fayyad 1996, p.5) visualises the concept of cyclical processes that form an interactive process by which data is refined into knowledge through sanitisation and contextualisation.

The data supplied from the Public Health system was in a pre-consolidated format. Specific directions with which to take the analysis presented themselves through each iteration, requiring various backsteps through the sequence as logically necessary. To guide these decisions systematically, a top-level framework for investigatory analysis needed to be established. Decisions were to be drawn from studies identified in the literature review and practical examples.

Data collection and quality

The Live Well Dorset programme's data were collected with the cooperation of data warehouse/analysis staff at Public Health Dorset. To initially inform various methodological decisions, a request was made for the entity relationship diagram (ERD) (Appendix A), this defined what data would become available and to how it was interrelated. Having anticipated the inclusion of various irrelevant data sets, this allowed for the development of an understanding as to what would be considered relevant to the research. To make the analysis process more efficient, it was used to identify what aspects of the system could be discarded. As a process this would be necessary when applying the same research approach to other organisations.

A subsequent meeting was conducted (21/06/2018) with database administrators and stakeholders to further interpret the available data and agree on the scope of the records to be extracted for analysis purposes. The resultant data structure can be seen below in FIG 5. The exploratory nature of the research approach concerning this area of the analysis meant a strategy which would gather as many fields that could be considered relationally relevant to each individual would be favourable. These include; physiological, categorical and demographic

information. Following the meeting, an export and data delivery strategy was agreed with the database manager who would complete the data delivery task. For convenience, and where it was deemed relevant, some data tables were also combined/flattened - a process that was understood not to have affected the integrity of the data. For interoperability between software systems, the data itself was exported in the commonly used data exchange format of comma-separated values (CSV).

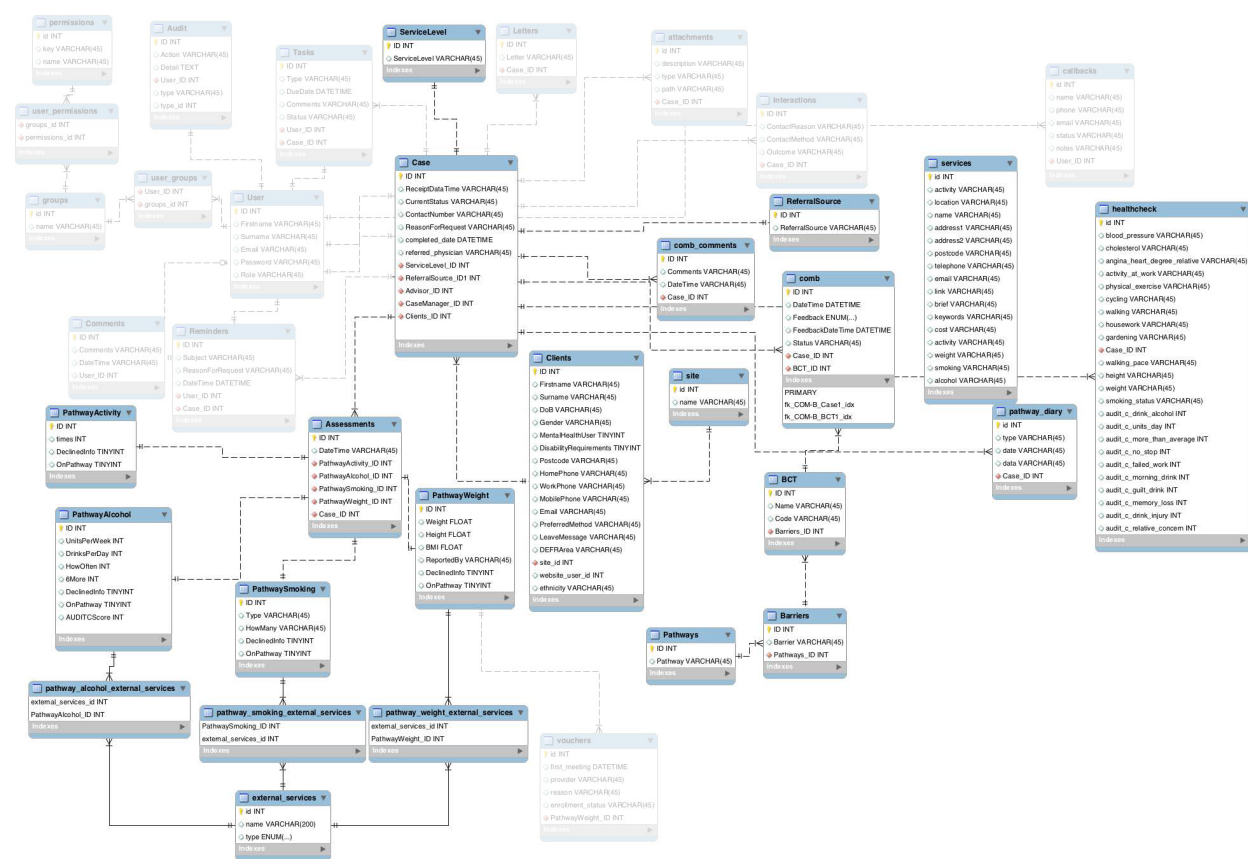


Figure 7. The database structural diagram from the Live Well Dorset programme where tables of information deemed irrelevant are greyed out and not considered for export. The resultant structure representing the complete data supplied for analysis can be found in Appendix B.

As an active system that is continually collecting data, the export process was also designed to be reusable. This was to ensure that further exports could be quickly and consistently repeated as more data became available over time. Specifically in the case of the researcher's efforts, time would pass between the development of the analytical tools and so the process was made easily repeatable.

As per the specification agreed with the Public Health Dorset data team, the export was supplied on (18/01/2019). Due to its content and privacy agreements with Public Health Dorset, the data cannot be distributed freely and instead must be explicitly requested from the organisation.

Sanitisation of the dataset was a key task to be performed before the computational analysis could begin. This involved cutting down to the fields within the tables of interest. Where duplicate rows were identified, the decision was to use the initial quantities for each pathway. This was straightforward apart from the smoking per day initial reading - the column appeared to contain non-numeric values, which were not in keeping with the expected data structure - investigation into mismatched columns in the source data was undertaken and resolved. Unexpected problems such as this is where the iterative nature of the data process became immediately applicable.

The process demonstrated a large potential for data inconsistency. Clients are not obligated to check in at each interval mentioned above (three, six and twelve months) after registration if they do not wish. Simultaneously, they are not motivationally bound to the service over time beyond their original intention to improve their wellbeing. If the service were not free of charge, a financial commitment would likely see greater retention. Gross (2001) disputes this assumption, pointing to their study of the same year (Gross et al, 2001), however, the results identify what could be considered numerous confounding variables, in part, due to a specific target demographic, which would not be applicable in a general study. This was not deemed relevant to the outcomes and has been disregarded. Therefore, there was an anticipation for irregular data consistency from the outset. Conversely, had a comprehensive survey (including the process of recruiting participants for the purpose of this study) been conducted, the volume of data obtained would be significantly lower than that of available from the LWD programme and not representation of the active system and processes.

Referral rate

Each registration included a range of data points (refer to Appendix B for a detailed list of all fields and their relational structure). As identified previously, the most relevant to the primary requirement of this research was the referral source and the timestamp captured at the point the client was entered into the system. The transformation of these two data points into a time series metric was used to satisfy the requirement for a 'GP referral rate'. The dataset time variable used for this purpose was the date on which the record was registered. The field was of date-time data type and the values were found to be accurate to the minute. A brief review of the data density was conducted to identify a suitable duration of the interval for analysis e.g., referrals per hour, per day, per week, per month. The expectation was that effects of communication would at very best take days to be seen, meaning that the decision was taken to group the data. This would bin referral counts by their calendar date (removing the minute of the day accuracy). As a result of this decision the referral rate dataset was established ready for use during further analysis.

Trend and distribution

To develop a general understanding, and by extension, confidence in the referral data, a series of distribution and trend analysis processes were proposed;

- Total referrals - a univariate understanding of the total referrals over the entire time period would identify a base to assess the relative magnitude of any records that may be identified and culled as erroneous in subsequent iterations of analysis.
- Distribution of records per month - in the context of a multi-year time series, monthly distribution of referral records was selected as a generalised method to assess the performance of the system.
- Referrals per month split by GP and non-GP (Appendix C) - splitting the record distribution by the focus variable to offer insight into its relative effect.
- Trend decomposition of the GP, non GP referral rate time series - identification of patterns brought about potential seasonal effects.

Approached as part of the knowledge discovery process, each analytical review was conducted to improve understanding of the data set in support of its purpose: causal inference of the effect of communication activities.

Demographics

An initial review of the supplied data structure (Appendix B) identified attributes suitable for demographic analysis of the client referrals. Understanding context would aid in the identification of potential confounders and biases, therefore improving the statistical confidence of any conclusions drawn as a result of the analysis. It was considered reasonable that stratification was to be used to identify areas of interest within the data. In an observational capacity, the first step was to explore data distribution and the baseline differences between groups. Specifically in the case of this research it was noted that within the case study data two groups were identified; those who were considered referred by a GP and those who were not.

3.3.4 Inferring causation

Where domain-specific studies have focused predominantly on isolated attributes of professional healthcare communication (see Chapter 2), they subsequently offer disparate insight when seeking to design and advance communication practices. This study set out to identify and quantify the potential effect of communication activities on the rate of GP referrals. Or, at least identify a correlation between intervention types and behavioural changes. Several approaches were considered which offered statistical significance testing in time series data. During the review of the GP and non-GP referral rates it was apparent that the separation of the data was relatively poor with some inconsistent distributions. If non-GP referrals were to be considered a control, and those referred by GPs to the treatment group, then a difference in difference approach would be applicable. However, the lack of clear separation of the two populations and awareness of the flexibility of the approach made causal inference a more desirable solution. This type of analysis models the counterfactual (what would have happened with no intervention) around a discrete point in the observed time series data. In the case of this study; the number of GP referrals.

Although not technologically trivial, the advantage of a causal inference solution is apparent when considering its relative simplicity in usage. However, its application and interpretation of the results will still depend on a given organisation's capabilities. For the researcher this was within capabilities, but in an on-going scenario for Public Health Dorset this may not have been the case.

An alternative approach using statistical change detection in time series was also considered. However, with the development of a chronology of known interventions, this process was deemed to be relatively redundant. If changes were identified and were not related to known events, it may have proved useful to consider their impact on the study, but was also considered an additional extension to the scope of an already large explanatory project. In summary, specific analytics looking for statistically significant rate changes that were unknown to the organisation was not taken further in this study.

The researcher's background in programming, as opposed to statistical mathematics, resulted in the decision to draw upon pre-existing analytical methods. Furthermore, the development of a new or novel statistical approach would have been beyond the scope of this study. The data analysis library selected for causal inference was Google Inc. *CausalImpact* (Brodersen et al, 2015) available for both the Python and R programming languages. The library was well established and its implantation appeared flexible and relatively abstracted from the underlying complexities of mathematics it facilitated. The researcher had little experience with either programming languages, but coming from a background in several other languages, Python was known to be the most syntactically similar to their existing expertise. Therefore, Python was used for the various data processing and analytical tasks detailed within this methodology.

3.5 Ethical Considerations & Risks

The use of public health data represented the primary ethical considerations in the context of this project. While the project was conducted independently, the project was commissioned by Public Health Dorset and the data which supported the research was managed and released within the organisation's ethical practices. In addition to this, the project was conducted within Bournemouth University's research compliance. As part of the institution's practices, the parameters of the project were assessed against the online ethics checklist tool (<https://ethics.bournemouth.ac.uk/>), and the subsequent review approved the research project (Appendix A). As a further step (although not required based upon the results of the University ethics review result because the data used was anonymous in nature and any results would not relate to any individual) the project was also reviewed against the Medical Research Council online tool, and was also shown to not require NHS REC approval.

More generally, the intended use of the data was towards the direct benefit of the public via the Live Well Dorset health and wellbeing service. The service places no restrictions upon requests to access the benefits, and is actively encouraged within the civic locality for which it is made

available. With the aim of the research to better enable, rather than diminish this activity, it did not propose to change the service itself, resulting in considering the activities undertaken to fall within the 'public good' (Ballantyne 2018).

Public Health Dorset excluded all data fields that contained personally identifiable data (name, address, date of birth). Furthermore, the one-way anonymisation of record IDs was undertaken to prevent reverse lookups. Randomly generated unique strings were inserted in place of the numeric IDs that represented individuals in the source data. Postcode locations were translated into the corresponding non specific LSOA geographic areas commonly used by the NHS (NHS 2018). These steps were deemed to be more than adequate in meeting data protection requirements (European Union 2016) while not impacting the meaningful attributes contained within the datasets.

The author did not foresee any noteworthy health and safety risks associated with the work to be undertaken in this research project. Digital data collection and analysis were anticipated to occur only at suitable locations and without the need of operation of unsafe equipment or exposure to unsafe practices affecting the individuals with which the study interacted.

The European Union's General Data Protection Regulation (2016) places legal obligations around the handling and use of personally identifiable data within EU member states. The process of anonymisation described previously as part of the ethical considerations absolved the data from the jurisdiction of these regulations; "...The principles of data protection should therefore not apply to anonymous information, namely information which does not relate to an identified or identifiable natural person or to personal data rendered anonymous in such a manner that the data subject is not or no longer identifiable." (European Union 2016).

3.6 Assumption

A range of assumptions were made and awareness was given to limitations that existed.

1. Many of the methodological decisions were directed by identifying relationships across broad processes, yet the data was specific to a geographic area and particular organisations. It was assumed that the results would be applicable to the case study organisation in question, and that this limitation is inherent of case study research. Operational and systemic comparisons would need to be researched to enable application to other public health organisations.
2. Of the referral data, random fluctuations were assumed to exist, therefore the identification of correlations between what would already exist as loosely related observations would be inconclusive if the variance was not sufficient enough to be probabilistically determined.

3. Where the chronological data were concerned, it was assumed that the request for information responses would be accurate and consistent enough to provide a meaningful historical record of interventions. More specifically, it was assumed that if tangible communication materials had been produced previously, copies would exist for review and classification by the researcher, and that the original distribution could be associated within a specific time period.
4. It was anticipated that demographic data would not impact the study in the form of confounding variables. This aspect of the data would be reviewed to measure the distributions for possible statistical anomalies, with the assumption being that if the demographic distribution was heterogeneous, then they would not be considered to represent confounding variables.

4. Results

4.1 Introduction

This chapter describes the results of the analysis conducted during this study. The chapter begins by describing the chronology of past communication interventions and the explorative case study dataset investigation. This is followed by the results from the cross-comparison analysis of the three data sources (intervention chronology, GP referral rate causal inference, and exploratory data analysis). Finally, the matrix developed as an extension to the literature review is defined, allowing for the classification of intervention communication activities.

Two parameters were required to satisfy the primary purpose of this research; definable communication activities with the intention of influencing GP referrals and Referral figures both before and after each communication activity. To increase confidence in the results of the primary analysis, exploratory investigation was undertaken to interrogate the demographic distributions of the referrals for abnormalities and patterns which could potentially bias and/or inform the primary analysis.

4.2 Intervention Chronology

The chronology shown below (Table 1) was intended to collect details of only those communication interventions which were explicitly targeted at GPs. During the process of collection it was discovered that very few specifically identifiable interventions in this capacity had taken place over the survey time period. The decision was made to extend the collection criteria to any events considered a potential influence on referrals to the LWD system. It was

proposed that this would increase the contextual understanding of referral rate analysis, and could facilitate the identification of false positives where both GP and non-GP events existed in close proximity.

Table 1. Chronology of Live Well Dorset service promotional communication initiatives

Identifier	Description	Date	GP Specific
A	'GP pack' a roughly 15 page document including posters, leaflets and cards hand delivered to each surgery by the outreach team. Details included pathway information and how to signpost / refer patients to the services.	2016	Yes
B	A one page document describing LWD and how to refer were laminated and distributed to surgeries with the intent they be on GP desks.	2017	Yes
C	Email to service stakeholders and partners about change of programme ownership from Optum to Public Health Dorset.	29th Mar 2018	No
D	New approach to social media. Personable, increased imagery.	30th Mar 2018	No
E	Updated website launched.	1st Apr 2018	No
F	Letter sent to GPs & Practice Managers from Dr Emer Forde - Local GP and Public Health Fellow.	1st Sep 2018	Yes
G	GP Email Bulletin via CCG.	1st Oct 2018	Yes
H	Supporting client tools 'My LiveWell' week long launch campaign.	18th Nov 2018	No

A coded list of communication intervention activities compiled from responses by supporting Public Health Dorset staff. Each item was to be supplied with as much detail available regarding the modality of each intervention, the period by which the intervention was operated and whether or not the intervention was specifically targeted at GP cohorts.

4.3 Referrals

4.3.1 GP and Non-GP Referral rates

The primary variable of this study to be analysed was the rate of referrals by GP sources. This would, by definition, rely upon the statistical analysis of the records in the form of a time series.

The focus of the research was to identify the effects on referral rates, in particular those referrals from GPs. This was to be achieved through analysis of referrals to the Live Well Dorset client relationship manager dataset. An initial review of the data showed there were 21,524 client pathway records within the data set for the time period 1st June 2015 to 12th January 2019. When initially identifying the records which were understood to be referred by a GP, some immediate inconsistencies became apparent. There were 115 unique categorical values used to identify the range of referral sources across two fields (due to a historic change in data structure), with multiple values a subset of which were regarded as 'GP' for the purposes of this study. To standardise the data, the decision was made to augment the records with a binary classification through identification of those values. In turn, this constituted a referral by a GP or by an associate GP practice staff. The inclusion of practice staff was made based upon the

diffusion of information within teams (Turner and Shepherd 1999) identified during the literature review as being an important aspect of GP communication within a practice setting. It was also assumed that this approach would maximise statistical density to the advantage of further analysis. The following referral classification values (Figure 6) were used to filter records which would be considered GP throughout the analysis that followed.

```
['Doctor',  
 'Doctor, Community Group',  
 'Doctor, Doctors',  
 'Practice Nurse',  
 'Health Care Assistant, Doctors',  
 'Midwife, Doctors',  
 'Nurse, Doctors',  
 'Practice Nurse, Doctors']
```

Figure 6. A programmatic output of original categorical naming conventions within the data set that were considered GP or GP influenced.

10,582 total records were identified as GP referrals after filtering had taken place.

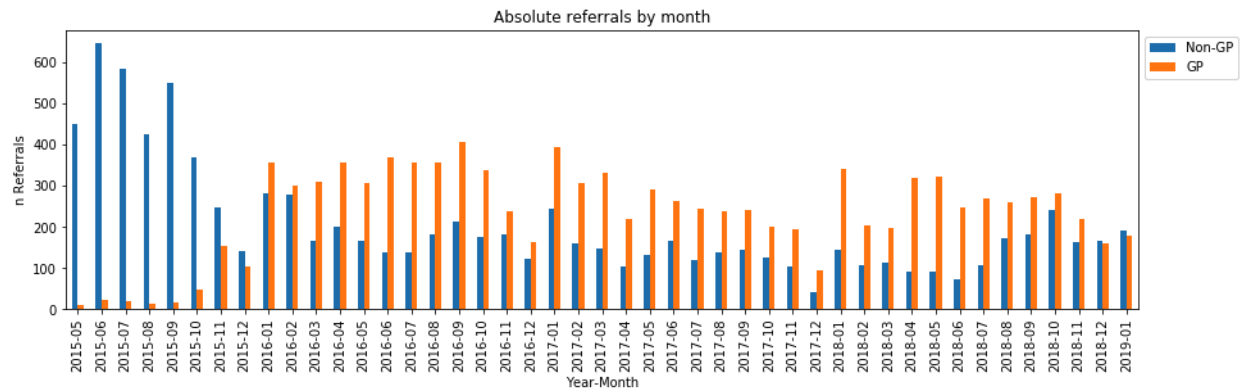
It was considered relevant to understand this figure in the wider context by identifying the number of records that had an unknown referral source. Of the 115 referral source values, two classifications were identified as constituting an 'unknown' source, specifically; "Other" and "Not asked". Records with these two values totaled 5,691. It was therefore possible to calculate GP referrals as a percentage of total known referral sources: 66.8%.

Registration distribution

Plotting total referrals by day was used to initially identify possible abnormalities in the distribution of referrals over time, the visualisation is more intuitively understood for time series data when compared to box plots or histograms. In this case it was also included to offer the researcher an overall appreciation of the density of data available to derive the results for the primary research problem. The count per-day referrals identified a significant spike on June 15th, 2016 (Appendix D), which was further identified as registrations taking place inside a two minute time period on this date. Through consultation with Public Health Dorset system staff this was described as "the CRM that our previous provider had was put in place in June 2016. The spike in data relates to historic weight loss data from our tier 2 weight management service, which needed to be uploaded into the new CRM." It was therefore considered important for the research to recognise that this import did not maintain the timestamps on which these referrals were originally gathered.

The potential impact on time series trend analysis was considered too pervasive to ignore. Therefore, the decision was made to remove the anomalous data from further analysis. This was completed by dropping all records within the times 17:03 and 17:04 on the date specified. This action resulted in a revised total of 10,512 GP referral records that would be used for subsequent analysis.

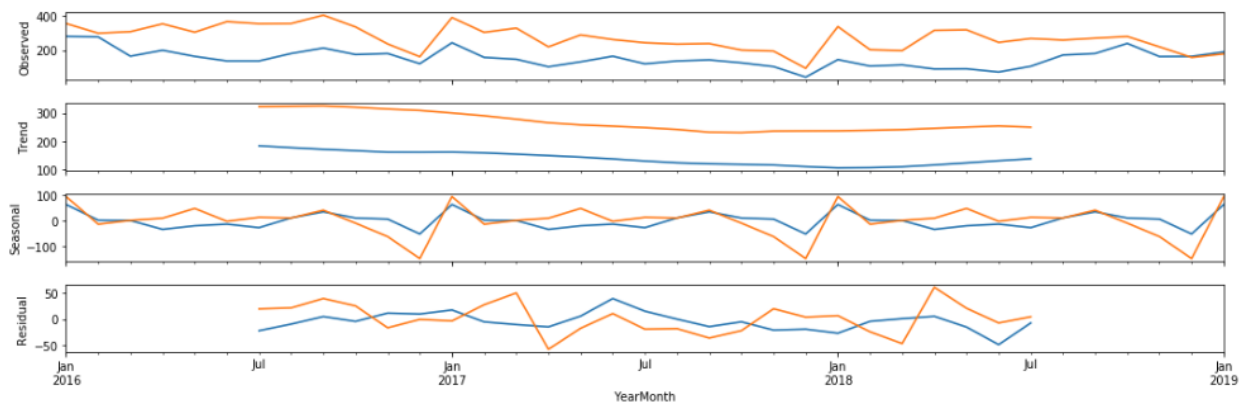
The distribution of GP and non GP referral counts by month (Graph 1) highlighted a large disparity in the time period for the months after the system went online. Aware that this period (up to and including December 2015) would impact interpretations of more recent (more relevant) trend analysis, the decision was made to remove these records from the dataset for all further time series analysis. This process resulted in 15,810 referral records, of which 10,120 were classified as from GP sources - a reduction of 3.73%.



Graph 1. Absolute GP and non-GP referral rates by calendar month. The separation of magnitude is clear in the visualisation. A distinct switch between a predominantly non GP to GP led referral rate taking place at the beginning of the time series was identified and subsequently removed, leaving a dataset starting January 2016.

Trend and seasonality

To understand underlying trends concerning the number of referrals over multiple years, a decomposition of the time series was performed using the *python statsmodels* library function *seasonal_decompose*. The results were then visualised in a paired axis plot of GP and non-GP referrals; see Graph 2.



Graph 2. Time series decomposition of the mean GP (orange) and non GP (blue) monthly referral rates. The graph was used to identify seasonal patterns that may (or may not) exist within the time series data. The process allows the researcher to explicate what may be relevant patterns and their influence on the potential recommendations of the study. Clear patterns can be seen with trough/peak pairings in the December/January date ranges, this is discussed in the following chapter (5.2.2).

Causal Inference

The causal inference analysis relied upon a distinct point in the time series being identified for each intervention in question. When no distinct date was available the dates analysed were themselves inferred from the results of the chronology gathering process. The researcher described the parameters that were used to standardise such circumstances as follows.

- Proximity - in cases where multiple events took place with less than an arbitrary seven day separation, interventions were combined and their midpoint selected as the intervention date for the purpose of analysis.
- Interval - interventions that spanned periods longer than one unit (a calendar day) were defined by selecting the midpoint of the time period identified in the chronology.

A 3-to-1 ratio of pre and post-intervention periods was applied. That is, the post intervention period was set to an arbitrary 60 day (2 month) interval, resulting in a 180 day (6 month) series used when modelling the post intervention counterfactual forecast.

For improved analysis, the non-GP time series was provided to the *CausalImpact* process as a control reference series. This posed a contradictory case for events which were not targeted at GPs but the decision to ensure comparison remained methodologically consistent for each intervention - the statistical implication of which is considered in the discussion.

The *CausalImpact* analysis library produced a range of statistical results for each intervention date;

1. *Predicted* - the counterfactual average number of daily referrals based upon the forecast from modeling the pre-intervention period. This value is a prediction of what would be expected, had no intervention taken place.
2. *Absolute effect* - the difference between the average observed and counterfactual (predicted) referral counts for the post intervention period.
3. *Relative effect* - a percentage difference between the average observed and counterfactual (predicted) referral counts for the post intervention period.
4. *Posterior probability p-value* - a value ≤ 0.05 suggests the null hypothesis is false, the assumption being the effect is unlikely to be a result of random fluctuations.
5. *Posterior probability of a causal effect* - percentage equivalent.

The results from analysis of each chronological intervention were collected and consolidated (Table 2). See Table 1 for details of each referenced intervention.

Table 2. Causal inference results of chronological communication interventions on GP referral rates

Intervention	Date	Actual	Causal Inference (Avg.)				
			Predicted (SD) [95% CI]	Absolute effect (SD) [95% CI]	Relative effect (SD) [95% CI]	Posterior prob. p-value	Posterior prob. of a causal effect
A	2nd Jun 2016	9.30	11.3 (1.5) [8.2, 14.2]	-2.0 (1.5) [-4.9, 1.1]	-17.5% (13.5%) [-43.0%, 10.0%]	0.099	90.11%
B	2nd Jun 2017	10.10	9.5 (1.1) [7.3, 11.6]	0.6 (1.1) [-1.5, 2.7]	6.4% (11.5%) [-16.1%, 28.8%]	0.296	70.43%
C, D, E	29th Mar 2018	8.00	5.7 (0.8) [4.2, 7.2]	2.3 (0.8) [0.8, 3.8]	40.0% (13.1%) [14.2%, 65.7%]	0.003	99.70%
F	15th Sep 2018	10.60	12.7 (1.1) [10.5, 14.7]	-2.1 (1.1) [-4.1, 0.1]	-16.3% (8.5%) [-32.5%, 0.7%]	0.032	96.80%
G	15th Oct 2018	10.80	12.7 (1.1) [10.5, 14.9]	-1.9 (1.1) [-4.1, 0.3]	-15.0% (8.9%) [-32.6%, 2.2%]	0.042	95.80%
H	18th Nov 2018	8.80	12.4 (1.1) [10.2, 14.6]	-3.6 (1.1) [-5.8, -1.4]	-28.9% (9.0%) [-46.9%, -11.5%]	0.001	99.90%

CausalImpact model periods (where available) were 180 days pre-intervention date, 60 day post-intervention date. A posteriori probability of >95% ($p \leq 0.05$) considers the null hypothesis to be false and a probable causal event/intervention effect having taken place on the referrals that followed.

4.3.2 Client Demographics

To investigate potential impact of referral demographics, relevant data was stratified across GP and non GP referrals. The results of which can be seen in Table 3. The relevance of the demographic data presented is discussed in the next chapter (5).

Table 3. Stratification of client demographics grouped by GP referral source

		GP Referral		
		isnull	No n = 10942	Yes n = 10582
Age, median [Q1,Q3]	New ros	54	52 [38,66]	52 [38,64]
Gender, n (%)	Female	0	8316 (76.0)	7788 (73.6)
	Male		2626 (24.0)	2794 (26.4)
Local Authority, n (%)	Bournemouth	0	2919 (26.7)	2757 (26.1)
	Christchurch		477 (4.4)	507 (4.8)

	East Devon		5 (0.0)	4 (0.0)
	East Dorset		980 (9.0)	1082 (10.2)
	North Dorset		699 (6.4)	844 (8.0)
	Not Available		408 (3.7)	298 (2.8)
	Poole		2442 (22.3)	2368 (22.4)
	Purbeck		675 (6.2)	578 (5.5)
	West Dorset		1058 (9.7)	1061 (10.0)
	Weymouth and Portland		1279 (11.7)	1083 (10.2)
Deprivation Quintiles, n (%)	20 to 40% most deprived	0	2329 (21.3)	2298 (21.7)
	20% most deprived		2980 (27.2)	2808 (26.5)
	40 to 60%		1918 (17.5)	1913 (18.1)
	60 to 80% least deprived		1954 (17.9)	1897 (17.9)
	80 to 100% least deprived		1353 (12.4)	1368 (12.9)
	Not Available		408 (3.7)	298 (2.8)

Stratification of focus data set demographics split by whether a referral was considered (by source categorisation) to be via a GP or non GP. "isnull" value included for data records that did not include an identifiable referral source.

4.3.3 Case Outcomes

To identify irregular distributions beyond those recognised in the limitations of this study, client-pathway interactions and subsequent health outcomes were analysed. The archival case records included detailed information about client biometrics and the associated progress through the wellbeing pathways with which they had registered. In total there were 22,793 case records for the period 1st June 2015 to 12th January 2019. With the purpose of the study surrounding referrals, the records were joined with their related client records to allow analysis of the case data to be conducted in context of the client, and importantly, their referral source.

An initial review of the data showed that the wellbeing pathway success variables did not appear to be consistent. As an example, it was noted that in the activity data, where a figure had been captured and no subsequent data existed, the status would default to a "Gone Down" which, without a measurement available, may or may not have been the case. In the interest of consistency and improved confidence (for this study), various algorithms to compute the pathway success parameters were developed (Appendix E). Beyond programmatic sanitisation logic, additional decisions were made which affect how the data was interpreted. The following list identifies the data sanitisation and potentially less intuitive parameters of the results presented in Table 4.

- A pathway is only considered active when both the activation flag and an initial measurement for that specific pathway biometric are present.
- A single record can have multiple pathway activations.

- Pathways which do not contain a second measurement at any of the 3 check-in periods are considered to have insufficient data and therefore neither a success or failure (isnull). The null value offers a measure of relative magnitude to those that complete a pathway.
- Smoking and Alcohol pathways require a reduction at the most recent check-in when compared to the initial measurement to qualify as a success.
- Weight loss pathways require a minimum of 5% reduction at the most recent check-in when compared to the initial measurement to qualify as a success.
- Activity pathways require any increase (measured in days) at the most recent check-in when compared to the initial measurement to qualify as a success.

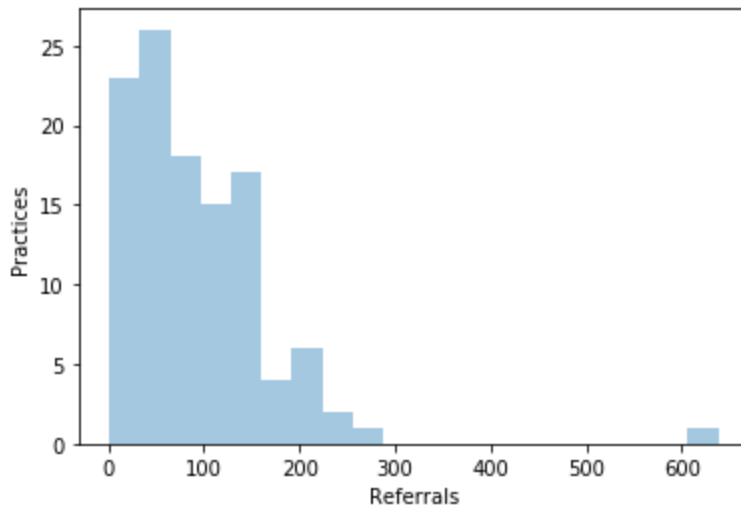
Table 4. Stratification of case pathway efficacy grouped by GP referral source

			Grouped by gp_referral	
			isnull	
			No n = 10762	Yes n = 10715
Registered Pathways, n (%)	0	0	907 (8)	818 (8)
	1		7724 (72)	7124 (66)
	2		1451 (13)	2065 (19)
	3		593 (6)	598 (6)
	4		87 (1)	110 (1)
No. of check-ins, n (%)	0	2050	6628 (69.2)	5031 (51.1)
	1		1730 (18.1)	2414 (24.5)
	2		1060 (11.1)	2028 (20.6)
	3		157 (1.6)	379 (3.8)
Reduced smoking, n (%)	-1	18086	233 (14.9)	128 (7.0)
	No		786 (50.3)	1464 (80.0)
	Yes		543 (34.8)	237 (13.0)
Reduced weight, n (%)	-1	5450	5941 (74.0)	4690 (58.7)
	No		997 (12.4)	1539 (19.2)
	Yes		1094 (13.6)	1766 (22.1)
Reduced alcohol, n (%)	-1	19646	732 (82.3)	683 (72.5)
	No		83 (9.3)	158 (16.8)
	Yes		74 (8.3)	101 (10.7)
Increased activity, n (%)	-1	17154	1119 (61.5)	1319 (52.7)
	No		551 (30.3)	901 (36.0)
	Yes		150 (8.2)	283 (11.3)

The pathway -1 denotes records with insufficient pathway measurements to classify the outcome

4.3.4 Practice referrals

The referral records identify 113 named GP practices that have referred patients. These accounted for 10,461 of the total referral figures analysed. The distribution of total referrals by practice was visualised in histogrammic form (Graph 3).



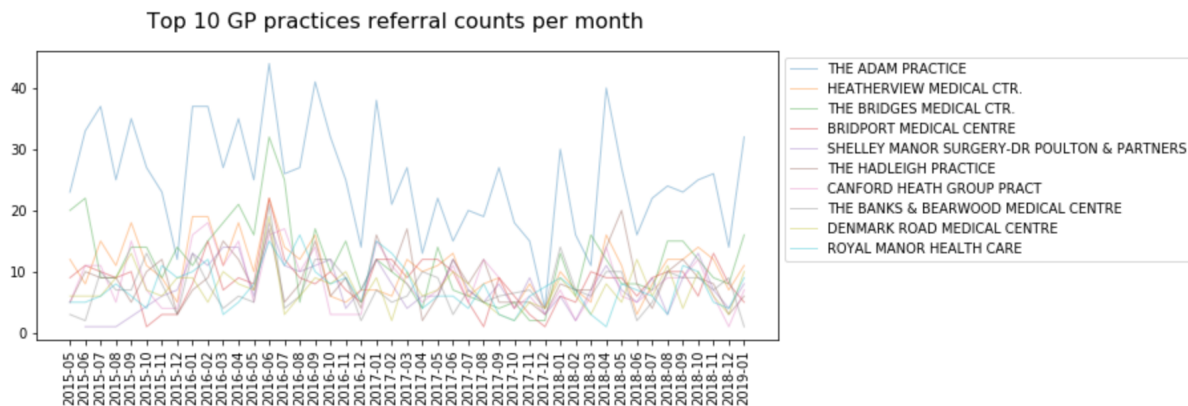
Graph 3. Histogrammic distribution of GP practice referral rates within the study dataset. Note the significant outlier at over 600 referrals detailed further in this chapter.

To extrapolate further understanding of the distribution of successful referral sources, and the outlier displayed in the histogram, the top 10 (by tidal referrals) practices were identified by name as shown in table 5 below. This process revealed “The Adam Practice” as the outlier outperforming the other high referring practices by more than double the total referral figures for the same period. This aspect of the results is discussed in more detail in the next chapter.

Table 5. 10 highest referring GP practices

GP Practice	Total Referrals
THE ADAM PRACTICE	638
HEATHERVIEW MEDICAL CTR.	279
THE BRIDGES MEDICAL CTR.	240
BRIDPORT MEDICAL CENTRE	232
SHELLEY MANOR SURGERY-DR POULTON & PARTNERS	221
THE HADLEIGH PRACTICE	218
CANFORD HEATH GROUP PRACT	217
THE BANKS & BEARWOOD MEDICAL CENTRE	198
DENMARK ROAD MEDICAL CENTRE	196
ROYAL MANOR HEALTH CARE	193

Due to the importance of time series analysis in the context of this study, the remaining step was concerned with forming a general understanding of how much greater (more desirable) referral counts were when distributed over time, and to identify any anomalous features among the top performing practices. To visualise the results, a plot of their competing referral activity over the complete time period was generated (Graph 4).



Graph 4. A line chart to visualise the comparative number of referrals the top 10 practices by total per month behaved in comparison to one another. The blue line “The Adam Practice” clearly stands out as having provided consistently more referrals.

5. Discussion

This study set out to identify the effect of communication activities on GP referral rates in a specific inter-professional healthcare setting. The research attempted to address the identification of measurable results within the data collected by Public Health Dorset as part of their service delivery. These insights were then to form the basis of a review of how effective the organisations strategies and process' in delivering improved performance of GP referrals to public health initiatives. The evidence presented in Chapter 4 has been formed through cross examination of chronological event intervention data and quantitative analysis of public health service records.

The well-being service data collected by PHD offered a range of statistical research opportunities, including the changes in referrals within this data and if these represented a relationship with the interventions. A method of causal inference was adopted to achieve this. This chapter begins by discussing the chronological intervention analysis and attempts to conceptualise the relevance in the wider context of the research. Each intervention is identified individually and discussed sequentially from oldest to newest.

This chapter also discusses the exploratory review of the original dataset. The purpose of this process was to attempt to identify any anomalous features that could impact the understanding of the causal results presented. In this supporting section (5.3) the researcher notes that this explanatory analytical process was intended only to extend contextual awareness/identification of the effect of potential confounders.

5.1 Chronology and causal inference

Generally, the quality of chronological data gathered was considered weak. In this context, good quality would be information about each intervention that accurately and fully satisfied each field of the proposed parameters detailed in section 3.3.1 of this document. While the content was suitably informative as a timeline - activities that took place historically were identified by staff - as the table of results (Table 3) displays, and interventions were recorded with wide interval periods and indeterminately tracked activities. This resulted in the need to define arbitrary, best-effort timestamp interpretations to satisfy the causal inference analysis requirements. It would, in the researcher's opinion, be prudent in future research to conduct and review intervention information before deciding upon the method of causal inference.

The causal inference analysis approach was best placed to determine the effect of discrete points in a time series. This made the arbitrary selection of mid-span time points a potential point of failure for accurate results. However, the resilience to incomplete or tumultuous time series data was nonetheless suitable for the task. Importantly, if the case-study organisation is considered operationally typical, the quality of the intervention tracking would not be unusual in

similar circumstances. This prompts the consideration of methodological changes to future studies, screening of archival intervention data before commencement of future studies, and promotion of improved tracking process at an organisational level.

Each intervention (see *Table 1* for related details) is considered in isolation before a summary is presented:

INTERVENTION A) Having selected a midpoint of a complete year of data (due to the lack of reliable intervention parameters provided), the interpretation of this intervention's results is considered highly speculative. The p-value is such that the marginal negative effect is considered to be spurious and not indicative of action undertaken around the point in time of intervention.

INTERVENTION B) The same midpoint approach as intervention (A) was applied to the only captured communication activity in the second calendar year of data collected, 2017. The marginal increase in average referrals between the counterfactual and observed data is, again, not considered statistically significant. However, in this particular case, the like-for-like date range allows a year-on-year comparison to also be made. A marginal increase (avg. 0.8 referrals) of the observed data is seen. Marginal growth trends can be expected as immeasurable communication interactions increase awareness of the service over time.

INTERVENTION C, D, E) This intervention encompassed 3 non-GP communication activities with their midpoint used for analysis. The span of only 3 days offered confidence in being a distinct intervention point. However, analysing GP referral data in the context of non-GP interventions presents a circumstance that is non-trivial when assessing the causal inference results. Here, a 40% increase in average number of referrals between the observed and the counterfactual is shown to be statistically significant with a low p-value of 0.003. The counterfactual, modeled upon the previous six months, sees a two-month post referral trend of average 2.3 fewer referrals than the observed values. Intuitively it would be reasonable to expect that non-GP interventions should not significantly affect the GP referral rate, which this result refutes. With this postulate in mind, the researcher considers possible confounders, which would present an alternate hypothesis:

- Misclassification - non-GP referrals categorised erroneously could account for an artificial increase unrelated to communication interventions. Such a phenomenon could occur as a result of changes in collection methods or data processes (pre-database). Additional correlation analysis should be sufficient to validate this hypothesis, followed up with consultation with case study system staff.
- Increases in public requests via their GP - increased public awareness which instigates requests via a GP could account for a potential increase from non-GP communication interventions. To validate this hypothesis, further analysis of the exposure of the communication initiative and greater detail pertaining to the demographic population would be required. To see a tangible increase appear, to the researcher, to be unlikely if

it is to be assumed people would not visit their GP specifically to access this service. This scenario would require a large enough subset of the original audience who require primary care during the two-month post-intervention period.

- Anomalous analysis - The *CausalImpact* library, in the case of this study, relies upon the default time series modelling, there is the possibility that in this case, the pre and post-intervention periods produce an anomalous counterfactual. To test this hypothesis, a customised model which fits the dataset more accurately could be substituted for causal analysis.
- GPs affected by public communication activity - outside of their vocational environments, GPs are themselves in the public domain and are therefore exposed to the same promotional materials as the public they treat. Interventions that are not GP targeted can also impact a GP audience. In the context of this study, the literature suggested the point that many communication practices can be considered common across both audiences.
- Unidentified activity - communication activities not captured as part of the chronology gathering exist as a potential cause of referral fluctuation. In the case of this study, and the recency of the intervention date, it is not believed that a significant enough communication activity would not have been included in the responses.

When considering these points in the context of RQ1, the ability to increase GP referral rates is a desired outcome, however, it can not meaningfully inform the research if the action taken is non-specific. Isolation of this one event as an extension to the research would be an obvious next step to further understand the outcome and what attribute of the intervention could be applied to further communication activities.

INTERVENTION F) This intervention was written content, from a peer source, and specifically targeted at the GP audience. These communication attributes meant it was considered particularly suitable in the context of the points identified in the literature review. The causal inference analysis presented a negative trend disparity between the referrals observed and the counterfactual (forecast) figure (average -2.1, -16.3%). Considering this alongside confidence interval and standard deviation, the result is not considered to be statistically significant. In the context of the communication activity undertaken, and it's positive attributes for informing GPs, it would also suggest that these figures are more likely the result of random fluctuations. Therefore, taking the causal inference result as accurate, while aware of the limitations (section 3.5), it is considered that this intervention activity was ineffective. As an alternate conclusion, it might be proposed that the time to observe an effect from this method of intervention is beyond the arbitrary two month post period counterfactual. In this case a further study would be required which took events proven to be effective and review the effect of variable pre and post period on causal inference of this type.

INTERVENTION G) A governing body targeted at the GP audience. In much the same way as intervention (F) was described, the causal inference reports a non-statistically significant result. The proximity of these two interventions means this result is not unexpected. Only a 30 day time

difference exists between them, so the counterfactual would be expected to be statistically similar, drawing on 66% of the same pre-period training data.

INTERVENTION H) The final intervention identified in the chronology was a public-targeted service delivery improvement. The referral dataset was insufficient to allow for comparable analysis on a two month post intervention period. As such, it was omitted from causal analysis.

As a whole, four GP target communication activities, two of which demonstrated highly variable time results from the chronological causal inference analysis, suggest ineffectiveness (in terms of statistically discernible changes) to the GP specific interventions to-date. With the method not being applied at its optimum, the questionable time parameters mean that this inference is offered with low confidence. The insights drawn are summarised in the context of the research project in chapter 6, Conclusion.

5.2 Referrals

As previously described, various processes were employed to separate GP referral sources from their non-GP counterparts (for code and explanation see Appendix D). The discussion that follows is not exhaustive, but placing the referral counts in the context of the data they are related to helps improve confidence. Applying the following domain knowledge also offers more understanding of how the activities surrounding the data collection and subsequent processing can be factored into the discussion of these results.

Continuous uncontrolled communication

During the collection of professional communication materials, it was identified that ongoing verbal promotion of the Live Well Dorset services is conducted in an ad hoc, unmeasured manner. It took the form of outreach teams who facilitate both the distribution of materials but also maintain unstructured interpersonal interactions with GP practice staff. This approach is therefore considered an uncontrolled variable of the study, and in-part guided the methodological decision to use the results on trend analysis which was hoped to be less susceptible to this type of ongoing communication efforts. When reviewing the results these sorts of activities should be considered as noise in the wider context - no record was supplied as to their proliferation geographically or over time.

Diffusion of information within interprofessional settings

Highlighted in the literature was the importance of GP to GP, or supporting practice staff to GP communication. An example scenario would be discussions of new processes such as the Live Well Dorset referral programme between GPs in external settings such as conferences, clinical commissioning group meetings, or materials being reviewed by practice staff and verbally communicated to GPs within the practice. These considerations led to the decision to include classifications that were not specifically GPs but do appear as though they would be indicative of these processes.

Misattribution

Self-referrals originally incentivised by GPs being classified as non-GP. Referrals at the point of capture are informed by human consideration and input accordingly. In this sense, they are interpreted and therefore open to error - a referral that may have been incentivised by a conversation with a GP could still be captured as a self referral. For example, if the individual considers the act of registration being the source, then they mark self-referred, rather than GP referred, when directed to registration by their doctor. Further to this point, in the time during the period in which the data has been collected and analysed, a web-based referral system was launched on the Live Well Dorset website and could be considered a further complication to source interpretation.

5.2.1 Absolute referrals by month

The absolute referrals by month (Graph 1) were presented to demonstrate the variation and comparative magnitude of the two primary variables of this study. Over the first six months (post May 2015 when data began to be collected), an obvious decline can be seen in the number of referrals until a December low point, followed by a significant uplift in January. Looking at the following two years (2016, 2017) offers some indication that this pattern is repeated, the same monthly period leading up to the new year demonstrates a sequential reduction. When considering an explanation as to this seasonal pattern, it is generally understood in the context of fitness and health that it is common for people to begin a new year with intentions of an improved lifestyle. Typically relating to health, the period leading up to this point sees these intentions deferred until the impending January 'new start' arrives. Looking at the most recent year, 2018, sees a lesser effect, and the preceding months see fewer total referrals. The lesser uplift in January can be attributed to an incomplete month of data as the final date in the dataset is known to be 06 January 2019. We would therefore anticipate a complete month to demonstrate the more regular pattern. Seasonality is discussed in more detail as part of the trend analysis section of this document.

The GP referral figures demonstrated a slow increase after the initial system implementation date. By not beginning high, a lesser drop is seen into December as the upward trend counteracts the seasonal reduction. The subsequent January (2016) figures display GP referrals overtaking the non-GP referrals, consistent for the following 33 months.

The increase observed towards the end of 2015 could be symptomatic of a slower uptake of promotional information. The seasonal trend gave some confidence in this observation - not seeing this pattern in the context of a service which supports typical lifestyle/new year goals would have introduced questions as to the validity of the audience.

As part of the referrals per practice review, the "Adam Practice" was highlighted for demonstrating noticeably more referrals than the next best performing surgery (638 vs 279

across the time series). This outlier was impossible to explain within the data collected, but a query to the Public Health Dorset staff resolved the point - the explanation simply that the practice was a noticeably larger practice (see section 5.3.5 for further discussion). This highlighted a shortcoming of the study, while efforts were made to check for confounders, not all confounders were known - in future research, note that collection of at least relative size of the practices should be undertaken. How different sized practices affect the diffusion of information offers an area for investigation, and could be achieved by analysing the growth in referrals compared to the number of GPs/supporting staff.

5.2.2 Referral Trend analysis

The seasonality component of referral rates in the time-series decomposition (Graph 2) both displayed patterns of note. The most obvious were apparent when reviewing the absolute values, the trough and peak in quick succession found pre and post respectively at the end of each calendar year. Through application of domain knowledge, it is proposed that this is resultant of less incentive for people to register for health and wellbeing services in the lead up to Christmas. Followed by the subsequent motivation to improve health and wellbeing at the start of a new year. This pattern offers confidence that the data is accurate because it represents real (expected) human behaviour. This observation also suggests it could be considered good practice to promote such services in such a way that these pre-existing patterns be utilised for increased service adoption - e.g. support GPs in referring patients in January, and avoid wasted resource use in December. Or, more accurately, prepare communicators and GPs with the knowledge in December that they will have better results with January referrals. It should be said that the transferability of approach will be largely societal, where populations exist that do not conform to this year-end cycle, it would offer little value. In such circumstances, seasonal composition should be run on a more applicable data set to inform similar decisions on efficacy of service promotion.

It was hoped that understanding not only the categorical source of the referrals but also the context of the practice by which the customer was associated would offer further insight into the efficacy of communication practices. As was clear in the results, one practice displayed a significantly larger referral rate than the other top performing locations (638 compared to 279). The initial reaction was that this provided a clear line of further investigation.

In consultation with Public Health it was identified that “The Adam Practice” is a particularly large GP surgery, and that this point alone is likely to account for the unusual outlier position it displayed within the referral figures. Not accounting for practice size in this aspect of the analysis was a shortcoming in the research design and will be highlighted in the conclusion of this study.

5.2.3 Demographics

Table 3 described the demographic distribution of client data. It was considered that there may be areas of investigation that would display differing demographic features between the GP and non-GP referral groupings. The consistency in numbers gives little opportunity to continue the line of enquiry.

While the results do not inform the primary concern of this study, considering the distributions outside of their referral source specificity shows regional distributions as expected when taking into account population density in these areas. An outlier exists in the form of East Devon, which is not within the official catchment of Dorset, so is assumed to be overspill where GP practices service their local catchment regardless of regional divisions. In all characteristics relative parity between referrals and non-referrals exists. No significant bias is seen in the delivery of referrals by GPs. The demographic results, therefore, show that the dataset is characteristic of a general distribution and should give confidence to the interoperability of further findings.

5.2.4 Case Outcomes

The case outcome results provided a top level understanding of how the service was being used. The purpose of this analysis was to see if any particular biases toward client types were seen, or results that may inform more understanding of the referral process - thus improving the ability of this study to explain the casual results obtained.

The obvious point of note was the large number of drop-offs after the initial registration/check-in. A free, no-committal service is anticipated to suffer from more casual usage, so is to be expected. Taking this as an indication of the clients being referred to the service, it would be possible to extend the research into measuring retention as a factor affecting the quality of referral. The value of improved retention could also offer more validity to statistical feedback (greater confidence and potentially more successful outcomes) when engaging GPs with referring their patients. Details of such considerations go beyond the scope of this study.

As a sense check this exercise was not considered particularly informative for the primary research question. It performed it's task, but the researcher would not be compelled to repeat this aspect of the analysis in repeated investigations unless a particular anomaly was identified in the data elsewhere that could be further informed by case outcome data.

6. Conclusion

This study investigated how GP rates are impacted in interprofessional healthcare settings. The research was conducted by means of a case study analysis of the Live Well Dorset program run by Public Health Dorset, who commissioned the study. Drawing upon their implementation and existing data sets to identify the efficacy of the organisation's efforts in this area. The aim of the study was to better understand and by extension inform future strategic and professional healthcare communication decisions. This chapter looks at the outcomes in response to the research questions defined in the introduction (section 1), the limitations that affected the study, and suggests directions in which to take further research.

6.1 RQ1 & RQ2

RQ1: In the context of Public Health Dorset, what (if any) effect can be identified as a result of previous communication interventions on GP referral rates to public health programmes?

The results of this study demonstrated that a time series analysis could be completed to measure the rates of referrals to public health programmes from GPs, in relation to the timing of a particular type of communication intervention. It is possible to use causal impact to compare GP referral rates both pre and post a communication intervention, to determine whether there was a statistically significant change in the number of referrals by GPs within a defined period shortly after the intervention.

Results revealed that communication events specifically targeted at GPs, A, B, F & G (see Table 1, Chapter 4 for details,) were not found to be associated with higher rates of GP referrals. Interestingly, it appears that some of their other strategies, interventions C, D, E & H, that were not specifically GP targeted did result in statistically significant increases in referral rates to the programs by GPs. Three of these four interventions, D, E & H, were largely targeted at the public, including a more personable social media approach, an updated website, and a launch of enhanced patient tools. This finding supports the idea that targeting the public is a more effective approach than targeting GPs. As discussed in section 1.2, GPs face high demands and time pressure, limiting consultation time with patients. Also, as identified by Peckham et al. (2015) and discussed in section 1.2, there has been questioning and debate by some GPs as to how much role they should play in public health initiatives. Taking these points into consideration, it is possible that GPs may not spend time discussing a referral to preventative health programs with a patient, unless initiated by the patient themselves - in other words, if there is a larger awareness within the public of these programs, patients may think to ask their GPs about them.

An alternative explanation to the noted higher rates of GP referrals within the time following a more publicly targeted approach, is that GPs are, of course, members of the public themselves.

It is plausible that the communication styles Public Health Dorset have used for their more publicly targeted initiatives may be simply more effective than the type of communication they are using to target GPs. If this were the case, then along with the rest of the public, a GP may respond more favourably to social media campaigns for example. In turn, this increased awareness could prompt them to refer their patients more. This point reiterates the attributes of effective communication discussed in the literature review, details of specific actions are compiled in the matrix found in section 2.2.4.

Intervention C involved a notice to stakeholders and partners about a change in ownership of the program from Optum to Public Health Dorset. This intervention was also associated with a statistically significant increase in GP referral rates. Unfortunately, this intervention fell in the same time period as interventions D & E above, therefore, no conclusions can be drawn as to whether the increased referral rates were actually truly associated with this intervention.

The study also revealed that there was an upswing in GP referrals consistently early in the year, annually in January, with a trough in December. This did not particularly correspond to a recognized “post period” of any intervention. This finding makes sense in the context of GP referrals being more patient led than GP led. As discussed in section 5.2.1, the start of a new year tends to be when many people begin working on new year's resolutions, often with a renewed focus on their health. It would make sense that they would approach their GPs, looking for help or advice on making positive lifestyle changes, such as quitting smoking, or losing weight. If interventions targeted toward the general public are more effective than GP targeted communication at increasing GP referral rates to public health programs, then it makes more sense for funding to be focused primarily on public initiatives.

RQ2: How effective was Public Health Dorset's strategy in improving referral rates among general practitioners?

It was evident from the causal inference results, that within the limitations of the data provided, further discussed below, there were no statistically significant impacts of GP targeted interventions by Public Health Dorset on GP referral rates to the Live Well Dorset programs. This leads to the conclusion that PHD's strategy of attempting to target GPs specifically has been ineffective.

There were several drawbacks to the strategy used that were believed to contribute to the effectiveness of intervention, many of which are discussed as drawbacks and limitations to the study. Data provided in the Live Well Dorset database revealed a large number of referral source categorisations, many of which were not clearly identifiable as GP or non-GP. For the purpose of this study, this meant that assumptions had to be made regarding which one of these two categories a referral source belonged. For example, a source from “doctor” was assumed to be a GP referral, as per table 6, section 4.3.1, although it was unknown what type of doctor this was referring to, specialist or generalist.

Additionally, the quality of intervention and chronological data itself was considered weak. Very few specifically identifiable intentional interventions had been recorded by Public Health Dorset with sufficient detail. The only intervention data available was that which appeared in Table 1. There was no framework or plan evident as to the type of communication method chosen as relates to the attributes of professional communication discussed in section 3 of this document. Interventions were recorded with wide interval times that were not well tracked, and time intervals of more than one intervention overlapped, making it very difficult to demonstrate relationships. There was also no data provided regarding the scope of audience of each intervention, so true exposure was not tracked - distribution locations and quantity of communication materials were unknown and unable to be accounted for in the analysis. This meant that rates of GP referrals from various clinics could not be correlated with the interventions specifically at each site.

The study results have important financial implications for Public Health Dorset, with an obvious potential for saving money by avoiding ineffective communication initiatives and practices. Further recommendations are discussed below in “future directions” to help improve future strategies toward communications regarding public health programming. An improved focus on data collection methods allowing more reliable research conclusions to be drawn, as well as using attributes of communication to help inform intervention decisions, are both key recommendations to help guide the most cost-effective approaches.

6.2 Limitations

- I. The study relied heavily on archival data sources, the accuracy of the data by definition limited the accuracy of results. It was necessary to assume that sufficient quality of data would exist such that the methodology proposed would enable the production of accurate analysis and subsequent results. Steps were taken to develop awareness of the underlying quality of data as described in Chapter 3; Quantifying Service Referrals.
- II. In the context of this study, communication interventions existed as the relationship between attributes of communication and the rate at which patients aka clients were referred to wellbeing services. This aspect was seen to heavily influence the effectiveness of the research, and would rely upon the quality of the historical intervention data gathered. Yet, the case-study organisation’s accuracy of recording activity in this area over time was outside of the control of the researcher. It therefore posed a significant limitation and was to be a key point to consider when reviewing results. As an extension to this, where the study saw time pass between development of the analytical tools and the point at which the results were produced, meant that the process of data delivery and the consistencies in collection processes (outside the control of the study) was assumed to remain consistent. It was therefore prudent to assume that in the period of data analysis that generally the collection methods employed by the case study organisation would also have remained consistent. For the

time series aspect of the data sources, in an effort to improve the analytical modelling and analysis, it was assumed that the use of data classified as 'non-GP' (referrals from sources other than those identified as GP based, see 4.3.1 of this document) would represent a sufficient control data set.

- III. The consistency of accurate referral data capture was considered outside the control of the study. Acquisition of real-world data offered more practically applicable results than under lab conditions, but was exposed to uncontrolled variations. The assumption was therefore made that in aggregate the process was to be considered consistent over time.

6.3 Future Directions

Within the context of the processes undertaken to gather information that allowed this study to progress, the following recommendations are proposed at an organisational level when designing communication activities:

1. Implement practices that capture (as a minimum) specific date, type, and quantity of audience exposure that each communication activity achieves. This activity should be conducted so as to not increase demand placed upon GPs (Riley 2018).
2. Identify suitable pre and post-periods - referral rate statistics should exist for the complete period.
3. Collect referral data for the post period. This approach is limited to analysis after the fact.

The first three recommendations, 1-3 above, are in response to the data itself, discussed in limitations (I) and (II). It is felt that following these would result in more accurate associations of interventions and their impact on referral rates, while allowing better identification of confounding factors. In a general sense, incorporating more detailed and completed data for the entire period of interest would allow a clearer understanding of true correlational relationships. As part of this suggestion, it would be an idea for GP referrals to record whether the patient themselves requested a referral from their GPs, or whether it was entirely initiated by the GP themselves.

4. Attempt to capture non-controllable variables that influence referral rates; e.g. the effect of collegial interactions within a practice (Dadich and Hosseinzadeh 2016).

Recommendation 5 recognizes the importance of uncontrolled variables as mentioned in limitation (III) and how they can impact actual data outcomes. The more that these variables are recognized and captured, the better the understanding of their true significance in relation to outcomes. For example, how effective GPs feel the particular public health programs are in actually changing their patients' clinical outcomes, which may come from conversations with colleagues, feedback from patients, or preconceived ideas.

5. Review each effective communication intervention against the attributes of professional medical communication as identified in Appendix A, and use this to inform future activities.

As a follow-up to this study and its clinical relevance, recommendation 6 is made to better inform more evidence based future decision making when creating type and method of communication interventions aimed toward health care providers, particularly GPs (Shamne and Nevzorova 2017; Horder et al, 1986).

Furthermore, where the public communication activities identified both GP and non-GP referral rate increases where non-gp interventions took place, it is proposed that further research would be applicable in understanding this effect. Specific analysis of this phenomenon could aid the identification of patient-driven GP referrals, which would reduce the relevance of professional medical communication in the context of this research and the case study organisation. This information could have an important impact in guiding a cost-effective approach to overall communications strategy.

It is proposed that adhering to the above activities, over a period of several communication interventions, would allow for continuous evaluation and optimisation of healthcare communications at an organisational scale. Capturing accurate logs of the results would also allow the potential for further studies to elaborate on the work conducted during similar research.

6. Study whether different types of incentives led to improved GP referral rates.

Although this study focused on communication type, and its effect on referral rates, as mentioned previously, there are several factors that impact referral rates. If communication in itself is not enough of a driver of needed referrals, it would be important to further study controllable variables such as specific incentives. As discussed by Peckham et al (2015), there has been a significant amount of distrust between the ever-changing public health organisations. Perhaps relying on altruism and simple knowledge about a program may not be enough to incentivise GPs to refer to programs, especially in the context of time constraints and multiple demands placed upon them.

6.4 Conclusion Summary

Overall, this study set out to identify the effect of various communication methods by Public Health Dorset on GP referral rates to their Live Well Dorset preventative health programs. The study reviewed the complexities of the GP role within the UK, particularly in relation to public health. The literature review also explored key attributes of effective communication, with a focus specifically on professional communication. A time series approach using causal inference was used to analyze data provided. The results revealed that a paucity of well defined and recorded data greatly limited the conclusions that could be drawn. Based on the data

provided, it appeared that the communication strategies targeted towards the general public actually had a greater impact on GP referral rates than those strategies targeted specifically towards GPs. There was also a noted upward trend in referrals early in each calendar year. It was recommended that further research be conducted to explore this area, as findings could have important implications for a cost-effective approach to public health program promotion. Further recommendations were made specifically regarding the importance of accurate and comprehensive data collection. Finally, the contribution of multiple other confounding variables that could impact GP referral rates were acknowledged, and noted to be important to consider in future research in this area.

References

- Akl, E.A. et al., 2011a. Framing of health information messages Cochrane Consumers and Communication Group, ed. *Cochrane Database of Systematic Reviews*, 22(1), pp.60–84.
- Akl, E.A. et al., 2011b. Using alternative statistical formats for presenting risks and risk reductions Cochrane Consumers and Communication Group, ed. *Cochrane Database of Systematic Reviews*, 58(8907), pp.455–90.
- Agarwal, N. et al., 2013. A Comparative Analysis of the Quality of Patient Education Materials From Medical Specialties. *JAMA Internal Medicine*, 173(13), pp.1257–1259.
- Ballantyne, A. & Schaefer, G.O., 2018. Consent and the ethical duty to participate in health data research. *Journal of Medical Ethics*, 44(6), pp.392–396.
- Baum, F., 2016. The new public health., (Ed. 4).
- Berendsen, A.J. et al., 2009. How do general practitioners and specialists value their mutual communication? A survey. *BMC Health Services Research*, 9(1), pp.1219–9.
- Boon, H. et al., 2007. Evaluating Complex Healthcare Systems: A Critique of Four Approaches. *Evidence-Based Complementary and Alternative Medicine*, 4(3), pp.279–285.
- Brodersen, K.H. et al., 2015. Inferring causal impact using Bayesian structural time-series models. *Annals of Applied Statistics*, 9, pp.247–274.
- Brotons C, Bjorkelund C, Bulc M, Ciurana R, Godycki-Cwirko M, Jurgova E, et al. Prevention and health promotion in clinical practice: the views of general practitioners in Europe. *Prev Med* 2005;40:595–601
- Bunge, M., Mühlhauser, I. & Steckelberg, A., 2010. What constitutes evidence-based patient information? Overview of discussed criteria. *Changing Patient Education*, 78(3), pp.316–328.
- Crawford, M.J., 2002. Systematic review of involving patients in the planning and development of health care. *BMJ*, 325(7375), pp.1263–1263.
- Dadich, A. & Hosseinzadeh, H., 2016. Communication channels to promote evidence-based practice: a survey of primary care clinicians to determine perceived effects. *Health Research Policy and Systems*, pp.1–12.

Denscombe, M., 2016. *The Good Research Guide: For Small-scale Social Research Projects*. Pp.1–389.

Department of Health. *The Mandate: A Mandate from the Government to the NHS Commissioning Board: April 2013 to March 2015*. London, Department of Health; 2012 Available from:

https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/256497/13-15_mandate.pdf

Dorset Clinical Commissioning Group, 2016. Update on sustainability and transformation plan (stp), with a focus on prevention at scale (pas). Available from: <https://www.dorsetccg.nhs.uk/wp-content/uploads/2019/03/08-Public-Health-Update-030419.pdf> [Accessed 09 July 2018].

European Union, 2016. *General Data Protection Regulation* [online]. Brussels: European Union. Available from: <https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:32016R0679&from=EN> [Accessed 09 July 2018].

Fayyad, U., Piatetsky-Shapiro, G. & Smyth, P., 1996. *From Data Mining to Knowledge Discovery in Databases*. Pp.1–18.

Goodwin, N., 2013. Understanding integrated care: a complex process, a fundamental principle. *International journal of integrated care*, 13, pp.e011–e011.

Goyder, C. et al., 2015. Email for clinical communication between healthcare professionals Cochrane Consumers and Communication Group, ed. *Cochrane Database of Systematic Reviews*, 54(3), pp.450–46.

Greengross, P., Grant, K. & Collini, E., 2001. *The history and development of the UK National Health Service 1948 - 1999*. Pp.1–39.

Gross, D. & Fogg, L., 2001. Clinical trials in the 21st century: The case for participant-centered research. 24(6), pp.530–539.

Gross, D., Julion, W. & Fogg, L., 2001. What Motivates Participation and Dropout Among Low-Income Urban Families of Color in a Prevention Intervention?*. *Family Relations*, 50(3), pp.246–254.

Hobbs, F.D.R. et al., 2016. Clinical workload in UK primary care: a retrospective analysis of 100 million consultations in England, 2007–14. *The Lancet*, 387(10035), pp.2323–2330.

Hogg W, Rowan M, Russell G, Geneau R, Muldoon L. Framework for primary care organizations: the importance of a structural domain. *Int J Qual Health Care* 2008;20:308–13

Horder, J., Bosanquet, N. & Stocking, B., 1986. Ways of influencing the behaviour of general practitioners. *J R Coll Gen Pract*, 36(292), p.517.

Irving, G. et al., 2017. International variations in primary care physician consultation time: a systematic review of 67 countries. *BMJ Open*, 7(10), p.e017902.

Johnson, A.E.W. et al., 2018. tableone: An open source Python package for producing summary statistics for research papers. *JAMIA Open*, 1(1), pp.26–31.

Knafllic, C.N., 2015. *Storytelling with data: A data visualization guide for business professionals*, John Wiley & Sons.

Knight, J. 2003. Scientific Literacy: Clear as Mud. *Nature* 423 (6938): 376–8.

Kothari, C.R., 2007. *Research Methodology : Methods and Techniques*. Pp.1–418.

Lalande, D. 2016. *Engagement and communications strategy: 2018-21 [online]*. Hounslow CCG. Available from:
<https://www.hounslowccg.nhs.uk/media/105799/091-Engagement-Comms-strategy-2018-21-GB-March-18.pdf> [Accessed 27 August 2018].

Leyland, A.H. & Groenewegen, P.P., 2016. Multilevel modelling and public health policy. *Scandinavian Journal of Public Health*, 31(4), pp.267–274.

MetaMetrics, 2019. The Lexile Analyzer [online]. MetaMetrics. Available from:
<https://lexile.com/educators/tools-to-support-reading-at-school/tools-to-determine-a-books-complexity/the-lexile-analyzer/> [Accessed 29 January 2019].

Mitchell H. Katz, 2011. *Multivariable analysis: a practical guide for clinicians and public health researchers*. Cambridge University Press.

Michie S, Atkins L, West R. (2014) *The Behaviour Change Wheel: A Guide to Designing Interventions*. London: Silverback Publishing.

Mogull, S.A., 2018. *Scientific and Medical Communication: A Guide for Effective Practice*, Taylor and Francis.

Nakai, M. & Ke, W., 2009. *Statistical Models for Longitudinal Data Analysis*. pp.1–11.

NHS, 2016. *About the NHS* [online]. UK National Health Service. Available from: <https://www.nhs.uk/using-the-nhs/about-the-nhs/the-nhs/> [Accessed 26 January 2019].

NHS, 2018. *Supporting Information: Lower Layer Super Output Area* [online]. UK National Health Service. Available from: https://www.datadictionary.nhs.uk/data_dictionary/nhs_business_definitions/l/lower_layer_super_output_area_de.asp [Accessed 20 September 2018].

NHS England, 2015. *Delivering the Forward View: NHS planning guidance 2016/17 – 2020/21* Available from: <https://www.england.nhs.uk/wp-content/uploads/2015/12/planning-guid-16-17-20-21.pdf> [Accessed 30 August 2018]

NHS Future Forum. Summary Report: Second Phase. London: Department of Health; 2012. Available from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/216422/dh_132085.pdf

NHS England, 2017. *GP Patient Survey 2017*. Available from: <https://www.england.nhs.uk/statistics/2017/07/06/gp-patient-survey-2017/> [Accessed 28 August 2018].

Onion, D., & Berrington, R., 1999. Comparisons of UK Family Practice and US Family Practice. *J Am Board Fam Pract.* 12(2)

Peckham S, Falconer J, Gillam S, Hann A, Kendall S, Nanchahal K, et al. The organisation and delivery of health improvement in general practice and primary care: a scoping study. *Health Serv Deliv Res* 2015;3(29)

Rice N, Smith PC. Strategic Resource Allocation and Funding Decisions. In Mossialos E, Dixon A, Figueras J, Kutzin J, editors. *Funding Health Care: Options for Europe*. Buckingham: Open University Press; 2002. pp. 250–71.

Riley, R. et al., 2018. What are the sources of stress and distress for general practitioners working in England? A qualitative study. *BMJ Open*, 8(1), p.e017361.

Rodrigues, A.M. & Sniehotta, F.F., 2018. A content analysis of the LiveWell Dorset service development phase: from behavioural diagnosis to intervention design optimisation. pp.1–40.

Rose, D., 2016. *Data Science: Create Teams that Ask the Right Questions and Deliver Real Value*, Apress.

The Kings Fund, 2017. The NHS: How providers are regulated and commissioned [online]. The King's Fund. Available from:
https://www.kingsfund.org.uk/sites/default/files/2017-10/NHS_structure_2017.pdf [Accessed 02 March 2019].

Schiavo, R., 2007. *Health communication: From theory to practice*, John Wiley & Sons.

Sergeev, D., 2018. Topic 9. Time series analysis in Python. Part 1. Basics [online].
<https://mlcourse.ai>. Available from:
https://nbviewer.jupyter.org/github/Yorko/mlcourse_open/blob/master/jupyter_english/topic09_time_series/topic9_part1_time_series_python.ipynb [Accessed 12 December 2018]

Shamne, N.L. & Nevzorova, M.S., 2017. Lingua-pragmatic potential of speech actions in medical professional discussions. In 7th International Scientific and Practical Conference "Current issues of linguistics and didactics: The interdisciplinary approach in humanities" (CILDIAH 2017). Paris, France: Atlantis Press, pp. 1–8.

Spiegelhalter, D., 2019. *The Art of Statistics: Learning from Data* (Pelican Books).

Sur, R.L. & Dahm, P., 2011. History of evidence-based medicine. *Indian journal of urology : IJU : journal of the Urological Society of India*, 27(4), pp.487–489.

Taylor, S.J. & Letham, B., 2017. Forecasting at scale. Pp.1–25.

Turner, G. & Shepherd, J., 1999. A method in search of a theory: peer education and health promotion. 14(2), pp.235–247.

Tversky, A. & Kahneman, D., 1981. The framing of decisions and the psychology of choice. *Science*, 211(4481), pp.453–458.

Van Bekkum, J.E. & Hilton, S., 2013. The challenges of communicating research evidence in practice: perspectives from UK health visitors and practice nurses. *BMC Nursing*, 12(1), pp.1–9.

Vermeir, P. et al., 2015. Communication in healthcare: a narrative review of the literature and practical recommendations. *International Journal of Clinical Practice*, 69(11), pp.1257–1267.

Wanless D. Securing Good Health for the Whole Population. London: HM Treasury; 2004
Available from:
<https://www.southampton.gov.uk/moderngov/documents/s19272/prevention-appx%201%20wanless%20summary.pdf>

Watson M. Going for gold: the health promoting general practice. *Qual Prim Care* 2008;16:177–85

Wyatt, J. & Sullivan, F., 2005. ABC of health informatics - What is health information? BMJ, 331, pp.566–8.

Appendixes

Appendix A



Research Ethics Checklist

About Your Checklist	
Ethics ID	21750
Status	Approved
Date Approved	21/01/2021 10:33:50
Date Submitted	20/01/2021 18:06:59
Risk	Low

Researcher Details	
Name	Andy White
Faculty	Faculty of Media & Communication
Status	Postgraduate Research (MRes, MPhil, PhD, DProf, EngD, EdD)
Course	Postgraduate Research - FMC
Have you received funding to support this research project?	Yes
Is this external funding?	Yes
RED ID	
Please provide the External Funding Body	
Is this internal funding?	
Please list any persons or institutions that you will be conducting joint research with, both internal to BU as well as external collaborators.	Public Health Dorset

Project Details	
Title	An exploratory case study analysis of the role evidence based communication plays in referral rates
Start Date of Project	01/05/2018
End Date of Project	01/05/2019
Proposed Start Date of Data Collection	01/06/2018
Original Supervisor	Einar Thorsen
Approver	Karl Rawstrone
Summary - no more than 600 words (including detail on background methodology, sample, outcomes, etc.)	

Review, through the use of statistical analysis of past, archival data collected by a local public health authority, the effect of distributed healthcare marketing materials have had on GP referral rates to public health services in interprofessional settings.

Filter Question: Does your study involve the use or re-use of data which will be obtained from a source other than directly from a Research Participant?

Additional Details	
Please describe the data, its source and how you are permitted to use it	The original data is collected, collated, filtered and processed by Public Health Dorset specifically for the purpose of analysis. In the case of this research project, a specific subset is provided which is relevant only to the analysis purposes described in the project description at a numerical/statistical level. This is the same process Public Health Dorset undertake for many of their analytical tasks with regard to the programmes they deliver for the local council authority and promote information in publicly accessible formats. The supply of the data is considered pre-approved and ethically considered through their own data handling adherences, and is understood to only be used to support the analysis within this research project.

Research Data	
Will identifiable personal information be collected, i.e. at an individualised level in a form that identifies or could enable identification of the participant?	No
Will research outputs include any identifiable personal information i.e. data at an individualised level in a form which identifies or could enable identification of the individual?	No

Storage, Access and Disposal of Research Data	
Where will your research data be stored and who will have access during and after the study has finished.	
The supplied data would be stored on the researcher's laptop. The laptop is passworded and only accessible to the researcher. Removal of the data will be undertaken upon completion of the research project.	
Once your project completes, will any anonymised research data be stored on BU's Online Research Data Repository "BORDaR"?	No
Please explain why you do not intend to deposit your research data on BORDaR? E.g. do you intend to deposit your research data in another data repository (discipline or funder specific)? If so, please provide details.	
The data is owned by Public Health Dorset, no agreement was made to distribute the data beyond the researcher and their research project.	

Dissemination Plans	
Will you inform participants of the results?	

Final Review	
Are there any other ethical considerations relating to your project which have not been covered above?	No

Risk Assessment	
Have you undertaken an appropriate Risk Assessment?	Yes

Data structure of archival dataset provided by the Live Well Dorset programme as at 16 May 2018.



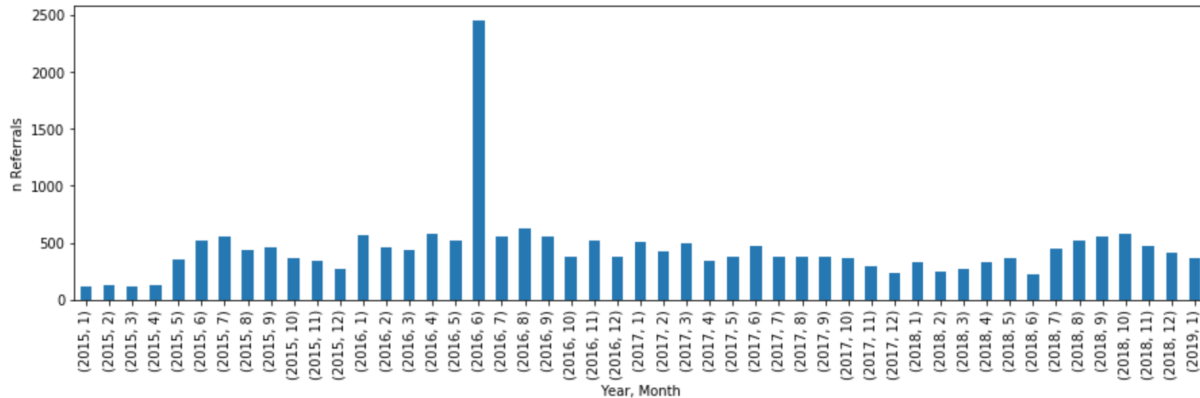
Appendix C

Raw (unfiltered) unique referral origin categories extracted from source dataset.

'Doctor' 'Event co-ordinator' 'Myself (Self-referral)' 'Friend' 'Nurse' 'Community worker' 'Pharmacy staff' 'Weight Loss Program' 'Manager at work' 'Health Visitor' "Children's centre staff" "Don't know" 'Midwife' 'Volunteer' 'Not asked' 'Exercise instructor' 'Health Care Assistant' 'Physiotherapist' 'Social worker' 'Carer' 'Occupational Therapist' 'Support Worker' 'Practice Nurse' 'Group leader' 'Doctor, Doctors' 'Event co-ordinator, LWD event' 'Nurse, Doctors' 'Nurse, Hospital' 'Community Health Worker, Community Group' 'Event co-ordinator, Health Check event' 'Nurse, Community Group' 'Doctor, Hospital' 'Pharmacy staff, Pharmacy' 'Weight Loss Program, Weight Watchers' 'Manager at work, Dorset County Council' 'Health Visitor, Doctors' "Children's centre staff, Centre" 'Manager at work, Hospital' 'Community Health Worker, Dorset & Wilts Fire and Rescue service' 'Weight Loss Program, Slimming World' 'Midwife, Hospital' 'Volunteer, Community Group' 'Doctor, Weymouth Comm Health Ctre' 'Friend, Weight Watchers' 'Nurse, Weymouth Comm Health Ctre' 'Myself (Self-referral), Doctors' 'Exercise instructor, Centre' 'Doctor, The Rosemary Health Ctr' 'Pharmacy staff, Well, 24 Crescent Street, Weymouth, DT4 7BX' 'Nurse, The Rosemary Health Ctr' 'Health Care Assistant, Doctors' 'Manager at work, Borough of Poole' 'Health Visitor, Canford Heath' 'Friend, Slimming World' 'Friend, Doctors' 'Myself (Self-referral), Social Media, Newspaper or Magazine Advert' 'Myself (Self-referral), Poster or Flyer' "Friend, Livewell's Website" 'Myself (Self-referral), Google Search' 'Health Visitor, Boscombe' 'Midwife, Dorchester' 'Midwife, Doctors' "Myself (Self-referral), Livewell's Website" 'Event co-ordinator, Active Ageing event' 'Doctor, The Panton Practice-Dr Sawyer & Partners' 'Community Health Worker, Health Check event' 'Physiotherapist, Hospital' 'Community Health Worker' "Myself (Self-referral), LiveWell's website" 'Myself (Self-referral), Social Media' 'Health Care Assistant, Hospital' 'Myself (Self-referral), Newspaper or magazine advert' 'Pharmacy staff, Doctors' 'Doctor, Cornwall Road Med Pract' 'Social worker, Dorset County Council' 'Myself (Self-referral), North Dorset' "Don't know, LWD event" 'Health Visitor, West Dorset' "Friend, LiveWell's website" 'Myself (Self-referral), Hospital' 'Myself (Self-referral), Dorchester' "Don't know, Doctors" 'Health Visitor, Purbeck' 'Doctor, The Marine & Oakridge Partnership' 'Weight Loss Program, Doctors' "Doctor, LiveWell's website" 'Doctor, Blandford St Mary, Blandford' 'Myself (Self-referral), Slimming World' 'Friend, Social Media' 'Occupational Therapist, Hospital' 'Doctor, The Poundbury Practice' 'Myself (Self-referral), Dorset County Council' 'Health Visitor, East Dorset' 'Health Visitor, Pokesdown' 'Health Visitor, Kinson, West Howe & Kingsleigh' 'Doctor, Community Group' 'Health Visitor, Dorchester' 'Friend, Hospital' 'Doctor, Townsend' 'Volunteer, Slimming World' 'Myself (Self-referral), Radio' 'Support Worker, West Dorset' "Don't know, LiveWell's website" 'Health Care Assistant, Slimming World' 'Wellbeing adviser' 'Community health worker' 'NHS health checks' 'Op Galaxy' 'Other' 'Wellbeing coach' 'Solutions4Health' 'Practice Nurse, Doctors' 'Group leader, Slimming World' 'Exercise instructor, Escape Pain Referral' 'Group leader, Weight Watchers'.

Appendix D

Distribution of referrals by month for complete data set time series including suspected false data.



Appendix E

The data sanitisation functions written in Python are included here as part of the wider analytical code produced as part of the result generation. These functions represent some of the more methodological decisions taken when classifying referral sources so were deemed particularly pertinent to the discussion within this document. The additional Jupyter notebooks can be accessed at the following URL <https://andyist.github.io/mres/>.

Live Well Dorset helper functions

Functions developed to specifically assist processing of Live Well Dorset wellbeing service data ready for further analysis.

```
# Load relevant libraries
import pandas as pd
import numpy as np
import math
```

1. Generate categorical values from client (demographic) data set

This function extrapolates categorical values for use during stratification and analysis. Results are returned as additional fields in the supplied pandas data frame.

- Gender (binary numeric)

- Month of year
- Age groups (10 year bins)
- Binary numeric GP referral source

```
def prepare_clients(df):
    # Numericise the gender field for arithmetic correlations
    df['gender_numeric'] = np.nan
    df['gender_numeric'][df.Gender == 'Male'] = 0
    df['gender_numeric'][df.Gender == 'Female'] = 1

    # Bin dates by month
    df['YearMonth'] = df.Date_registered_Month_Year.map(lambda x: x.strftime('%Y-%m'))

    # Bin ages by 10 year divisions
    groups = ['0-9', '10-19', '20-29', '30-39', '40-49', '50-59', '60-69', '70-79', '80-89', '90+']
    df['age_group'] = pd.cut(df.Age, range(0, 101, 10), right=False, labels=groups)

    # Mark GP referrals by practice staff and those with "Doctor" in the referral source identifier that are not hospital related
    ref_filter = ['GP', 'Doctor', 'Doctor, Community Group', 'Doctor, Doctors', 'Practice Nurse', 'Health Care Assistant, Doctors', 'Midwife, Doctors', 'Nurse, Doctors', 'Practice Nurse, Doctors']
    df['gp_referral'] = 0
    df['gp_referral'].loc[df['ReferralSource'].isin(ref_filter) | df['how_a'].isin(ref_filter) | df['how_a'].isin(ref_filter) | df['referral_combined'].isin(ref_filter)] = 1
    df['gp_referral'].sum()

    # Mark GP referrals by practice staff and those with "Doctor" in the referral source identifier that are not hospital related
    ref_filter = ['Doctor', 'Doctor, Community Group', 'Doctor, Doctors', 'Practice Nurse', 'Health Care Assistant, Doctors', 'Midwife, Doctors', 'Nurse, Doctors', 'Practice Nurse, Doctors']
    df['gp_referral'] = 0
    df['gp_referral'].loc[df['how_a'].isin(ref_filter) | df['referral_combined'].isin(ref_filter)] = 1

    # Dataframe of unknown referral sources
    #ref_filter = ['Not asked', 'Other']
    #unkown_refs = df.loc[df['how_a'].isin(ref_filter) & df['referral_combined'].isin(ref_filter)]

    return df
```

2. Output a list of all referral types

This function returns a list of the unique referral sources found within the data set.

```
def referral_types(df, fields=['how_a', 'referral_combined']):
```



```
return pd.unique(df[fields].values.ravel('K'))
```

3. Count the number of unique referral types

This function includes the number of records found for each of the unique referral source classifications

```
def get_unique_referral_types(df):  
    return pd.unique(df[['ReferralSource', 'how_a', 'referral_combined']].values.ravel('K'))
```

4. Smoking quantity standardisation

This function compares multiple related columns of varying data types and decides to what level the client is/was a smoker and translates this to a simplified numerical categorisation.

- 0 - Non smoker
- 1 - Smokes 1 to 9 times a day
- 2 - Smokes 10 to 19 times a day
- 3 - Smokes 20 or more times a day

```
def smoking_class(values):  
    # Define standardise codes for various data values  
    mapping = { 'a. 1-9': 1, 'a.': 1, 'b. 10-19': 2, 'c. 20+': 3, 'New non smoker': 0, 'Non-Smoker': 0, 'Non  
smoker': 0, 'Cigarettes': 1, 'e-Cigarettes': 1, 'Roll-ups': 1, 'Smoking (cigarettes, cigars, pipe, roll-ups):  
1, 'E-cigs / vape (E-cigarettes)': 1 }  
    missing_flag = False  
    zero_flag = False  
    coded = False  
    for value in values:  
        if value in ['Missing']:  
            missing_flag = True  
        if value in ['0', 0]:  
            zero_flag = True  
        if value in mapping:  
            coded = mapping[value]  
            continue  
        if not coded:  
            if value.isdigit():  
                value = pd.to_numeric(value)  
                if value >= 1 and value <= 9:  
                    coded = 1  
                elif value < 19:  
                    coded = 2
```

```

        elif value >= 20:
            coded = 3
    if not coded or missing_flag and zero_flag:
        return np.nan
    else:
        return coded

```

6. Standardise case pathway data

This function reviews and consolidates tracking data for the 4 service pathways offered by Live Well Dorset. Through initial analysis it was found that numerous aspects of the data were incomplete or poorly structured (multiple interpretable values). The formalised variables are appended to the original data frame for use in further analysis. Definitions of each pathway's 'success' state are detailed in the accompanying thesis document.

```

def prepare_cases(df):
    # Loop all case records
    for index, row in df.iterrows():
        pathway_count = 0
        followup_count = []
        pathways_success = np.NaN
        # Smoking reduction pathway
        if row['Smoking_PathwayActivatedFlag'] == 1:
            pathway_count += 1
            # Ignore rows who start as non smokers (this should be captured with the above, but better to be sure)
            if (row['Smok_Init_Qty_Day_Group'] != '0') & (row['Smoking_InitialQtyPerDay'] != 'Non smoker') & (row['Smoking_InitialQtyPerDay'] != '0'):
                smoking_followups = 0
                smoking_success = 0
                # Standardise the smoking quantities
                smoking_df = {}
                start_smoking =
                smoking_class(row[['Smoking_InitialQtyPerDay', 'Smok_Init_Qty_Day_Group']].values)
                smoking_df['Smoker3'] =
                smoking_class(row[['Smoking_ThreeMthFUpQtyPerDay', 'Smoking_3_Month_Daily_Quantity']].values)
                smoking_df['Smoker6'] =
                smoking_class(row[['Smoking_SixMthFUpQtyPerDay', 'Smoking_6_Month_Daily_Quantity']].values)
                smoking_df['Smoker12'] =
                smoking_class(row[['Smoking_TwelveMthFUpQtyPerDay', 'Smok_12_Months_Qty_Day_group']].values)
                # Process each potential followup
                for key in smoking_df:

```

```

    if not math.isnan(smoking_df[key]):
        smoking_followups += 1
        end_smoking = smoking_df[key]
    # Flag if smoking has reduced
    net_smoking_reduction = float('nan')
    if end_smoking < start_smoking:
        smoking_success = 1
        pathways_success = 1
        smoking_net_reduction = start_smoking - end_smoking
    else:
        if smoking_followups > 0:
            smoking_success = 0
        else:
            smoking_success = -1
    df.at[index, 'smoking_followups'] = smoking_followups
    followup_count.append(smoking_followups)
    df.at[index, 'smoking_success'] = smoking_success

# Weight loss pathway
if row['Weight_PathwayActivatedFlag'] == 1 and row['Weight_Initial'] > 0:
    pathway_count += 1
    weight_followups = 0
    weight_success = 0
    end_weight = row['Weight_Initial']
    for field in
['Weight_ThreeMthFUpWeight', 'Weight_SixMthFUpWeight', 'Weight_TwelveMthFUpWeight']:
        # Process each potential followup
        if row[field] > 0:
            weight_followups += 1
            end_weight = row[field]
        # Flag if weight has been lost
        net_weight_loss = float('nan')
        if end_weight < row['Weight_Initial']:
            weight_net_loss = row['Weight_Initial'] - end_weight
            weight_net_loss_percent = (row['Weight_Initial'] - end_weight) * (100/row['Weight_Initial'])
            if weight_net_loss_percent > 5: # Success is a 5% loss or more
                weight_success = 1
                pathways_success = 1
            else:
                weight_success = 0
        else:
            if weight_followups > 0:
                weight_success = 0
            else:
                weight_success = -1

```

```

df.at[index,'weight_followups'] = weight_followups
followup_count.append(weight_followups)
df.at[index,'weight_success'] = weight_success
df.at[index,'net_weight_loss'] = net_weight_loss

# Alcohol reduction pathway
# Ignore rows who start as non drinkers (this should be captured with the above, but better to
be sure)
if row['Alcohol_PathwayActivatedFlag'] == 1 and row['Alcohol_InitialAlcoholUnit'] != 0:
    pathway_count += 1
    alcohol_followups = 0
    alcohol_success = 0
    end_alcohol = row['Alcohol_InitialAlcoholUnit']
    for field in
['Alcohol_ThreeMthFUpAlcoholUnit','Alcohol_SixMthFUpAlcoholUnit','Alcohol_TwelveMthFUpAlcohol
Unit']:
        # Process each potential followup
        if row[field] != -1:
            alcohol_followups += 1
            end_alcohol = row[field]
        # Flag is alcohol has reduced
        net_alcohol_reduction = float('nan')
        if end_alcohol < row['Alcohol_InitialAlcoholUnit']: # This is not units, they have to have
dropped a banding to be 'successful'
            alcohol_success = 1
            pathways_success = 1
        else:
            pathways_success = 0
            # alcohol_net_reduction = row['Alcohol_InitialAlcoholUnit'] - end_alcohol This is a
categorised value - not calculable
        df.at[index,'alcohol_followups'] = alcohol_followups
        followup_count.append(alcohol_followups)
        if alcohol_followups > 0:
            df.at[index,'alcohol_success'] = alcohol_success
        else:
            # If there were no follow ups it is unfair to say the pathway did not succeed
            df.at[index,'alcohol_success'] = -1
        df.at[index,'net_alcohol_reduction'] = net_alcohol_reduction

# Increased activity pathway
if row['Activity_PathwayActivatedFlag'] == 1 and row['Activity_InitialActivityLevel'] > -1:
    pathway_count += 1
    activity_followups = 0
    activity_success = 0
    end_activity = row['Activity_InitialActivityLevel']

```

```

    for field in
['Activity_ThreeMthFUpActivityLevel','Activity_SixMthFUpActivityLevel','Activity_TwelveMthFUpActivit
yLevel']:
    # Process each potential followup
    if row[field] != -1:
        activity_followups += 1
        end_activity = row[field]
    # Flag is activity has increased
    net_activity_increase = float('nan')
    if end_activity > row['Activity_InitialActivityLevel']: # This is not units, they have to have
dropped a banding to be 'successful'
        activity_success = 1
        pathways_success = 1
    else:
        pathways_success = 0
        # alcohol_net_reduction = row['Alcohol_InitialAlcoholUnit'] - end_alcohol This is a
categorised value - not calculable
    df.at[index,'activity_followups'] = activity_followups
    followup_count.append(activity_followups)
    if activity_followups > 0:
        df.at[index,'activity_success'] = activity_success
    else:
        # If there were no follow ups it is unfair to say the pathway did not succeed
        df.at[index,'activity_success'] = -1
    df.at[index,'net_activity_increate'] = net_activity_increase

    df.at[index,'pathways_success'] = pathways_success
    df.at[index,'pathway_count'] = pathway_count
    if len(followup_count):
        df.at[index,'followup_count'] = max(followup_count)
    else:
        followup_count = 0
    return df

```

6. Client-Case join

This function joins client records to the related cases.

```

def join_clients(df, clients):
    # Merge clients with cases data set
    df = pd.merge(df, clients, left_on='clientID', right_index=True)
    # Remove duplicate case IDs
    df.drop_duplicates(subset="CaseID", keep='first', inplace=True)
    return df

```

7. Format time periods for Causal Impact analysis

With a given daily time series dataframe and a specified intervention date, this function returns a pre and post from/to value pair for use with the causal impact library. Intervals (in days) can be optionally included to override the defaults.

```
def get_periods(df, intervention, post_interval = 28, pre_multiplier = 3):
    post_from = min(item for item in df.index if item > intervention)
    post_to = intervention + timedelta(days=post_interval)
    try:
        post_to = min(item for item in df.index if item > post_to)
    except ValueError:
        # We are past the end of the dataset, use the max and inform instead
        post_to = max(df.index)
        print("Warning; max series date used for post_to: " + str(post_to) + ". Post period = " + str(post_to
- post_from))
    pre_from = intervention - timedelta(days=(post_interval * pre_multiplier))
    pre_from = min(item for item in df.index if item > pre_from)
    # Get date immediately before intervention
    loc = df.index.get_loc(intervention)
    pre_to = df.index[loc]
    return [str(pre_from.date()),str(pre_to.date())], [str(post_from.date()),str(post_to.date())];
```

8. Generate table data outside of Jupyter notebook

This function is to assist with outputting tabular data for use in other formats.

```
def output_html_table_file(mytable):
    data, metadata = get_ipython().display_formatter.format(mytable)
    with open('tables/referral-demographic-stratification.html', 'w') as f:
        f.write(data["text/html"]) # Assuming the object has an HTML representation
```